The Alberta Journal of Educational Research

Vol. V, No. 1 *March, 1959*

THE COMMITTEE ON EDUCATIONAL RESEARCH

Faculty of Education
University of Alberta

ACKNOWLEDGMENT

This publication was made possible by funds granted by Carneige Corporation of New York. That Corporation is not, however, the author, owner, publisher, or proprietor of this publication, and is not to be understood as approving by virtue of its grant any of the statements made or views expressed therein.

The Alberta Journal of Educational Research

THE COMMITTEE ON EDUCATIONAL RESEARCH
Faculty of Education
University of Alberta

VOL. V, No. 1 MARCH, 1959

Editor: H. E. Smith

CONTENTS

PAGE

EDITORIAL 5

The Many Voices: A Preliminary Investigation Into the Consistency of the Self Concept 7
C. C. Anderson

The Relationship Between Socio-Economic Status, Intelligence and Attainment in Some Scottish Primary Schools 16
W. B. Dockrell

Alberta Composite High Schools and Gifted Youth 23
I. E. Housego and G. L. Mowat

An Introductory Study of the Consistency of Marking Mathematics Papers in Schools 31
D. H. Crawford

Achievement and Under-Achievement in English Language Ten in an Alberta Composite High School 41
R. S. Chapman

Teacher Education in England 50
H. S. Baker

JOURNAL REVIEW: G. M. Dunlop 64

BOOK REVIEW: Ann Dockrell 66

BOOK REVIEW: W. B. Dockrell 68

NOTICE OF PUBLICATION: Monograph 72

Editorial

Research in education is many-sided: statistical, comparative, historical and descriptive. Each has its place, with much overlapping of implications.

Our schools—their organization, curricula, methodologies, preparation of teachers, and even the school buildings—are the products of historical processes which invite assessment and understanding. Features of programing, standards, and cultural environments are topics for comparison, as Baker has demonstrated in the December issue and in this.

Papers like those of Anderson (September, 1958) and Crawford (this issue) pose questions of interpretation. The manifest tendency of pupils to answer test items somewhat differently on repeated tests, and the obvious disposition of teachers to assess the same papers differently on different occasions, call for answers on freedom of choice. Perhaps the best one can do is at the descriptive level in calling attention to the fact and extent of variation, and suggesting measures of precaution and control. Perhaps one should extend the research to repeated choices in value preferences as in Art or Literature, as well as in factual preferences as of knowledge items.

This sort of research might (but probably would not) throw light on the age-old philosophical question of freedom of choice vs. determinism in human responses. Determinism is in fact implied in the expectation of consistency of response to any given stimulus. Inconsistencies or variations likewise are assumed explainable on the grounds of different states of the organism, such as post-prandial sluggishness, fatigue, or variant mental states of euphoria, remorse, rebelliousness and the like.

Similarly a book review such as that of Dockrell in this issue (A Study of Interpersonal Relations) suggests questions of board interest. How much are people prone to faddishness in the espousal of theories? In a recent address, Dr. MacPhearson, Superintendent of the Oliver Mental Hospital, recalled some fads that have been hailed as sure-cures, or cure-alls, of abnormal mental states: mesmerism, suggestion, psycho-analysis, insulin, electric shock, inter-personalism, the open door. All are potentially useful as therapeutic means, but none constitutes a massive breakthrough—a phrase too often heard in educational research circles. Mr. Dockrell does not speak as a faddist.

Education and psychiatry have much in common. Both aim to overcome the ill effects of much childhood training or lack of it.

Both attempt to remove road blocks to learning in the form of repressions, inhibitions and false beliefs. Both attempt to adapt training programmes to individual differences. Both aim, at least in part, at well-integrated, functioning personalities. And finally the factors are so involved and complicated that massive breakthroughs are highly unlikely. In both cases it seems desirable that research should be conceived broadly, historically as well as statistically, philosophically as well as descriptively.

In this issue Baker treats English teacher-training institutions descriptively, historically and comparatively. Crawford examines statistically and speculatively teacher consistency in marking mathematical papers. Chapman attempts to isolate statistically the factors associated with under-achievement in English language. Justification for the provision of kindergarten training is deduced by Dockrell from his mathematical analyses of relationships among socio-economic status and various quotients. And finally, Housego and Mowat assess, largely through personal interviews, the extent to which special talent is recognized and capitalized upon in typical Alberta composite high schools.

THE MANY VOICES: A PRELIMINARY INVESTIGATION INTO THE CONSISTENCY OF THE SELF CONCEPT

CHARLES C. ANDERSON
Faculty of Education
University of Alberta

The fall and rise of the self concept is now a well-worn theme in historical psychology. Self-concept theories, as morally respectable substitutes for psychoanalysis, enjoy considerable vogue at present and critics, excluding Hall and Lindzey (8, p. 498), keep at a respectable distance, and even Hall and Lindzey based their critical attacks on counter-assertions of a qualitative type. The present paper, by contrast, presents a preliminary experimental examination of two fundamental tenets of all self-concept theories. In the course of this, the self concept is taken with Bills (4) to mean "the traits and values which the individual has accepted as definitions of himself." (5, p. 257), and the ego-ideal is regarded as the self the individual is striving to become.

The main tenet to be examined is the consistency of the self concept. This notion was prominent in Horney (9) and Lecky (13) who saw the striving for self-consistency as a primitive need. In Lecky's terms, "The nucleus of the system, around which the rest of the system revolves, is the individual's idea or conception of himself. Any idea entering the system which is inconsistent with the individual's conception of himself cannot be assimilated, but gives rise to an inconsistency which must be removed as promptly as possible." (13, p. 145)

Later exponents of this theme have included Rogers (18) Jersild (11 and 12) and Smith, Bruner and White (19, p. 33). Indeed Jersild extends the area of self-consistency to include the ego-ideal which he tends to confuse, at least syntactically, with the self.[1] "In his endeavour to maintain self-consistency, a person is likely not only to seek and select experiences that are in keeping with the image of himself that he *would like to maintain* but also to resist and reject experiences that go counter to this image." (12, p. 355) The validity of the application of the consistency hypothesis both to the self concept and to the ego-ideal is open to experimental investigation. This investigation was carried out in the present instance at the grade 8 level because, although self-estimates might be expected to fluctuate as a result of the quick changes in mood so characteristic of adolescents and preadolescents, Hurlock

[1]This confusion is evidently the product of Jersild's following Horney's view that the ego-ideal, consisting of idealized neurotic strategies, is perceived by the individual as his "genuine" self.

(10) has stated that "This concept of self is established in the early years of life, taking its fundamental form in the years of babyhood and childhood. While it unquestionably does change to a certain extent, the changes are, for the most part, slight." (10, p. 468)

The second tenet to be examined is one originating with G. H. Mead: that the individual's attitude to himself is determined largely by the early attitudes of "significant others" to him. If this is true, then the extent and direction of the fluctuation or inconsistency of the self concept between various testing occasions should be in phase with the extent and direction of the fluctuation or inconsistency in the individual's interpretation of group attitudes towards him. If the two sets of inconsistency estimates are not correlated and do not covary, then it would seem unlikely that the individual's self concept is largely a product of the attitude of "significant others". This question is also open to experimental investigation, although it should be pointed out that, under the time-interval conditions of the present experiment, it is assumed that the self concept changes within a fairly short time (say a few days) after any encounter with different attitudes on the part of others.

This investigation into the validity of the consistency has one important implication for a prominent theoretical position adopted by self-theorists. This is the view often expressed by Jersild and Rogers that the root of anxiety and maladjustment lies in the felt discrepancy between an individual's self concept and his ego-ideal or idealized version of himself. The accuracy of such a theory will be impaired if significant amounts of fluctuation are observed between testing occasions in either the self concept or ego-ideal or both. At the moment some authors, notably Calvin (5), Phillips (17), McIntyre (15), Berger (3), and Omwake (16), report research findings which support the view that there is a relationship between the self concept and the attitudes of others to the self, while other investigators, notably Levy (14), Cowen (6), Zimmer (24) and Zuckerman (25) report evidence which indicates no such relationship. Perhaps this discrepancy is a product of self-inconsistency.

Method

Test Instrument

The present research, then, is a preliminary experimental investigation into the consistency of measures of the self concept, the ego-ideal and group attitudes towards the self, and into the covariation among the set of inconsistency indices available for these three variables or functions. The relatively simple testing instrument was constructed from items used by Taschuk (21) in his Self Concept Inventory and was divided into three parts. The self-

concept test consisted of twenty items such as : "I am an agreeable person." The tests of group attitudes towards the self consisted of the same twenty items rephrased so that the subject of each sentence was always "friends" or "people who know me well" or "friendly acquaintances", and the object was always the pupil: for example, "Friends think I am an agreeable person." For the self-ideal test, the twenty items were again rephrased to take the form "I should like to . . .": for example, "I should like to be an agreeable person." The twenty items in the self-concept test were divided into parallel forms by constructing half of the items to be synonymous with the other half. Thus "I am an agreeable person" was matched with "I am a pleasant person", and "I am self-reliant" with "I have confidence in myself." Exactly the same parallel matching of A with B forms was made in the other two tests so that for example, "I should like to be a self-reliant person" was matched with "I should like to have confidence in myself." The pupil was asked to rate on a five-point scale the extent to which each of the 60 items or statements was applicable to, or true of, himself. The 60 items, arranged in a random order, were administered as a group test on four different testing occasions, each separated by an interval of one week, to 105 pupils in grade 8, 47 boys and 58 girls, with an average I.Q. on the Terman-McNemar Test of Mental Ability of 100.3, the standard deviation being 13.8.

Method of Statistical Analysis

The method of testing for the existence and extent of function fluctuation or inconsitency was that devised by Thouless (23) and recently applied by the present writer. On the self-concept function, for example, each pupil made A and B scores on each of four testing occasions (a total of eight scores), and this was also true of the other two functions. For each function separately the eight sets of scores were intercorrelated and Thouless's test was applied to the pattern of intercorrelations between every combination of two testing occasions in the following way. Take for example the first two occasions on which the self-concept test was administered; each individual has four scores, A, and B (parallel form administered on the first testing occasion) and A_2 and B_2 (parallel forms administered on the second testing occasion). The intercorrelations comprise the within-occasion coefficients ($r_{A_1B_1}$ and $r_{A_2B_2}$), the between-occasions coefficients ($r_{A_1B_2}$ and $r_{B_1A_2}$) and the test-retest coefficients ($r_{A_1A_2}$ and $r_{B_1B_2}$). Thouless's criterion for evidence of fluctuation is the correlation of differences $r_{(A_1-A_2)\ (B_1-B_2)}$, and the index[2] for estimating the extent of fluctuation is usually taken as

[2]The rationale behind the derivation of both criterion and index is given by Thouless (1936).

the criterion divided by the average of the within-occasion coefficients. For each of the three functions independently, self concept (S), ego-ideal (E), and group attitude (G), there were available six possible combinations of two testing occasions, and consequently six criteria and indices.

Results

Consistency of the Functions Tested

The average of the six criteria for each of the three functions is presented in Table I.

TABLE I
AVERAGE FOR EACH FUNCTION OF SIX CRITERIA FOR EVIDENCE OF FLUCTUATION

	Functions		
	S	G	E
Average	.284**	.118	.181

Note—**Significant at the 1% level.

There is clear evidence here of inconsistency or fluctuation in the self concept as it is tested in this way. However in the amount of observed fluctuation the three functions do not differ significantly, as Table II shows.

TABLE II
EXTENT OF FLUCTUATION BETWEEN FOUR TESTING OCCASIONS OF EACH OF THREE FUNCTIONS INDEPENDENTLY

Pairs of Occasions	Indices		
	S	G	E
1 and 2	1.227	—2.693	1.419
1 and 3	1.133	2.000	.314
1 and 4	2.327	— .848	.142
2 and 3	—1.194	2.500	1.586
2 and 4	1.641	.061	.370
3 and 4	1.846	.798	.149

The test for the significance of differences among the mean indices is made in Table III.

TABLE III

THE SIGINFICANCE OF DIFFERENCES AMONG GROUP MEANS OF DATA FROM TABLE II

Source of Variation	df	Mean Square	F
Between Groups	2	49,243	.68
Within Groups	15	72.126	.....

It is clear that the means do not differ significantly among themselves and that the functions underlying S. G. and E scores fluctuate to approximately the same extent. That this extent is serious can be shown by comparing any of these three average indices with the average index of fluctuation for a formally similar Annoyances Test previously administered by the writer (1) to a group of slightly younger British children. The average index for the Annoyances Test, which showed clear evidence of fluctuation, was only .415 or about one third of the average indices reported in the present investigation. This suggests that, at this level of testing, there are many selves, and, in particular, scores on the self-concept and ego-ideal tests are so inconsistent that any attempt to measure "self-acceptance" by subtracting one score from the other is bound to lead to findings which are very variable among different researches.

Testing Mead's Hypothesis

The second fundamental tenet to be examined is the positive and causal relationship said to exist between the self concept and group attitudes towards the self. This is not merely a matter of estimating the static correlation between scores on the S and G tests because, as F. V. Smith (20) has indicated, evidence of a causal relationship depends not on coincidence but on covariation. If, then, group attitudes have early determined the nature of the self concept, we might expect that any inconsistency in scoring on the G test would covary both in extent and direction with any inconsistency in scoring on the S test. The first procedure, therefore, was to combine the A and B scores of each individual for each function (S, G and E) and for every testing occasion separately, so that he had three (A + B) scores for S, G and E tests on each of four testing occasions. In particular each individual had four S scores (S_1, S_2, S_3 and S_4) and four corresponding G scores (G_1, G_2, G_3 and G_4) and, isolating for a moment the first two testing occasions, we see that Thouless's criterion in the form $r_{(S_1-S_2)\ (G_1-G_2)}$ will yield evidence

of the desired covariant association between inconsistency indices on the S and G tests. Since there are six possible groups of two testing occasions, there will be six criteria and indices for each of the three groups of two functions, S and G, E and G, and S and E. The relevant criteria for S and G are presented in Table 4.

TABLE IV
CRITERIA FOR EVIDENCE OF RELATIONSHIP BETWEEN FLUCTUATION IN S AND G FUNCTIONS

Pairs of Occasions	Criteria
1 and 2	—.084
1 and 3	.323**
1 and 4	.265**
2 and 3	—.185
2 and 4	.696**
3 and 4	.215*

Note—*Significant at the 5% level.

The average correlation here is .235, which is significant at the 5% but not at the 1% level. However before we may adduce this as complete evidence in support of Mead's position, it is in order to indicate that the significant covariation between the inconsistency indices of the two functions does not tell us which function is cause, which effect. When Smith (20) talked of a causal explanation being implied in a covariant relationship, he had in mind the covariation, at a simple level, of the presence or absence of an introduced experimental factor with the presence or absence of some observed effect. Such a situation does not obtain in the present investigation and, to determine the validity of Mead's position, we must have recourse to deducing that, if that position is accurate, then the extent of the indices available for the other two combinations of functions (S and E, and E and G) must be significantly smaller than the extent of the indices observed for the S and G combinations. This will tend to be true of the S and G combinations because presumably the ego-ideal will be less causatively determined than the self concept by the G factor, and it will be more markedly true of the S and E combination which is not under the hypothetical influence of the G factor at all. The raw material for testing this hypothesis is presented in Table V.

TABLE V

EXTENT OF FLUCTUATION BETWEEN FOUR TESTING OCCASIONS IN THREE COMBINATIONS OF PAIRED OVERLAPPING FUNCTIONS

Pairs of Occasions	Indices for Combinations of Functions		
	S and G	E and G	S and E
1 and 2	— .144	.637	—1.214
1 and 3	.708	2.161	1.462
1 and 4	.366	1.213	3.963
2 and 3	— .395	—1.879	.940
2 and 4	1.009	— .093	1.297
3 and 4	.352	1.179	—1.469

The test for the significance of the differences among the group means is presented in Table VI.

TABLE VI

THE SIGNIFICANCE OF DIFFERENCES AMONG GROUP MEANS OF DATA IN TABLE V

Source of Variation	df	Mean Square	F
Between Groups	2	.455	.16
Within Groups	15	2.840	.....

It is clear that these means do not differ significantly among themselves and that consequently these data do not substantiate the hypothesis that the self concept is determined by group attitudes to the self. This suggests that the three functions tested interact with and reinforce each other in a more complicated fashion than Mead imagined.

Discussion

The self-report made at the paper-and-pencil level shows evidence of many selves. This view does not go unsubstantiated by observers of human behaviour working at a rather deeper level. As has been noted by Anderson (1), the inconsistency view has inspired a characteristic Maugham short story. La Rochefoucauld (Fitzgibbon, 7) frequently noted that nothing is more changeable than man (axiom 65), that self-interest speaks many tongues and plays many parts (axiom 39) and that "At times a man is as dif-

ferent from himself as he is from others." (axiom 135, 7, p. 58) And Thomas (22, p. 26) in a poem headed The Letter, talks of

". . . what the hand has written
To the many voices' quiet dictation."

Summary

Tests of the self concept, the ego-ideal and of group attitudes towards the self were administered to 105 grade 8 pupils. There was evidence of considerable fluctuation or inconsistency in the functions being measured by these tests, and there was no evidence to support Mead's contention that an individual's self concept is determined very largely by the attitude of "significant others" to the self.

REFERENCES

(1) Anderson, C. C. Function fluctuation. *Brit. J. Psychol. Monogr.*, 1958, No. 30.

(2) Anderson, C. C. The extent and factorial structure of function fluctuation in a group of adolescents and preadolescents. *Brit. J. Educ. Psychol.*, to appear.

(3) Berger, E. M. The relation between expressed acceptance of self and others. *J. abnorm. soc. Psychol.*, 1952, 47, 778-782.

(4) Bills, R. E., Vance, E. L., & McLean, O. S. An index of adjustment and values. *J. consult. Psychol.*, 1951, 15, 257-216.

(5) Calvin, A. D. Adjustment and discrepancy between self-concept and inferred self. *J. consult. Psychol.*, 1953, 17, 39-44.

(6) Cowen, E. L. The negative self-concept as a personality measure. *J. consult. Psychol.*, 1954, 18, 138-142.

(7) Fitzgibbon, C. (Ed.) *The maxims of the Duc de la Rochefoucauld.* London: Wingate, 1957.

(8) Hall, C. S., & Lindzey, G. *Theories of personality.* New York: Wiley, 1957.

(9) Horney, K. *New ways in psychoanalysis.* New York: Norton, 1939.

(10) Hurlock, E. B. *Child development.* New York: McGraw-Hill, 1953.

(11) Jersild, A. T. *In search of self.* New York: Teachers Coll., Columbia Univ., Bureau of Publications, 1952.

(12) Jersild, A. T. *The psychology of adolescence.* New York: Macmillan, 1957.

(13) Lecky, P. *Self-consistency.* New York: Island Press, 1945.

(14) Levy, L. H. The meaning and generality of perceived actual-ideal discrepancies. *J. consult. Psychol.*, 1956, 20, 396-398.

(15) McIntyre, C. J. Acceptance by others and its relation to acceptance of self and others. *J. abnorm. soc. Psychol.*, 1952, 47, 624-625.

(16) Omwake, K. I. The relation between acceptance of self and acceptance of others, shown by three personality inventories. *J. consult. Psychol.*, 1954, 98, 443-446.

(17) Phillips, E. L. Attitudes towards self and others: a brief questionnaire report. *J. consult. Psychol.*, 1951, 15, 79-81.

(18) Rogers, Carl R. *Client-centered therapy.* Boston: Houghton Mifflin, 1951.

(19) Smith, M. B., Bruner, J. S., & White R. W. *Opinions and personality.* New York: Wiley, 1956.

(20) Smith, F. V. *The explanation of human behaviour.* London: Constable, 1949.

(21) Taschuk, W. A. An analysis of self-concepts of grade 9 students. *Alberta Journal of Educational Research,* 1957, 3, 94-103.

(22) Thomas, R. S. *Poetry for supper.* London: Faber, 1958.

(23) Thouless, R. H. Function fluctuation and test unreliability. *Brit. J. Psychol.,* 1936, 26, 325-343.

(24) Zimmer, H. Self-acceptance and its relation to conflict. *J. consult. Psychol.,* 1954, 18, 447-449.

(25) Zuckerman, M., Baer, M., & Monashkin, I. Acceptance of self, parents and people in parents and normals. *J. clin. Psychol.,* 1956, 12, 327-332.

THE RELATIONSHIP BETWEEN SOCIO-ECONOMIC STATUS, INTELLIGENCE AND ATTAINMENT IN SOME SCOTTISH PRIMARY SCHOOLS

W. B. Dockrell
Faculty of Education
University of Alberta

Frequent investigations have been made into the relationship between socio-economic status and intelligence test score. Binet (1) himself noted the relationship and many investigations have been made since in England and America. Duff and Thomson (2) were among the earliest British investigators. Terman and Merrill (10) reported their findings when they standaridized the Stanford-Binet in America. The correlations were positive but not high—Duff and Thomson .28, Terman and Merril .40. When the Social criterion takes account of the cultural and educational level of the home as well as the economic level correlations are slightly higher e.g. Leahy (7). 51. Some investigations, for example those of Neff (8) in America and Lawrence (6) in England have been concerned with the influence of the environment. Most investigators have reported (i) a correlation between socio-economic status and I.Q. and (ii) the possibility of the influence of environmental as well as hereditary factors.

The Problem

This study was undertaken to examine further the relationship between Intelligence Test score, attainment and socio-economic status. The observed relationship alone tells little. The most casual inspection of school results will show that children of a higher socio-economic group tend to do better at school than those lower down the social scale. It may be argued that intelligence is hereditary and that the superior intelligence of children of a higher socio-economic group is inherited from parents themselves of superior intelligence. On the other hand higher test scores may be a function of better schooling, better food, rest and recreation, better cultural facilities in the home or greater parental interest. It is possible that the large differences in attainment between social groups are due in part to better schooling, but investigations using Intelligence tests reduce this advantage, and Fleming (3) has shown that in Scotland schooling is not the most important factor.

To elucidate the relationship two hypotheses were formulated:

(1) that the correlation between I.Q. and socio-economic status would increase between the beginning and the end of primary

schooling because environmental factors influencing the correlation would continue to operate, and

(2) that the correlation between socio-economic status and intelligence can wholly or partly be explained by differences in acquired skills; these differences in acquired skills being a result of different environmental influences.

The Method

The first hypothesis was tested by calculating correlation and regression coefficients between socio-economic status and change in I.Q. between ages 7 and 11; the second by analysis of test quotients on standardized tests of Intelligence, English and Arithmetic at age 11 and the relationship between these quotients and socio-economic status.

Previous investigations have used very different assessments of socio-economic status though most were based on parental occupation or objective ratings. This investigation was conducted in a Scottish city where the Local Education Authority was unwilling to allow an individual assessment of socio-economic status. However, an assessment prepared by School Welfare Officers was made available. This gave the percentage of the population of each school area in each of five social gradings. These percentage ratings were converted into a single composite estimate of socio-economic status giving gradings from 0 to 27.

For the first part of the investigation the Intelligence Quotients of 1,062 children from 30 Schools on Moray House Intelligence Tests for 7 year olds and Moray House Intelligence Tests for 11 year olds were used. Both these tests had been given as part of the schools' normal testing programme. The schools were selected to give a typical sample of the distribution of socio-economic status. School mean quotients were calculated for both tests and correlated with socio-economic status. Regression and Correlation Coefficients were calculated between social status and change in I.Q.

The second part of the investigation included a Factorial Analysis and the calculation of partial correlation coefficients of school mean social gradings and scores on Moray House English, Arithmetic and Intelligence tests of all children completing primary school in the city in 1952. Again the tests were all part of the normal programme.

The Results

As far as possible large schools were chosen for the first part of the investigation so that chance differences could cancel each other out. In one school where there were only 6 pupils the mean decline

in I.Q. for five of them was 3.75 points of I.Q. but a decline of 20 I.Q. points for the sixth child increased the mean change in I.Q. for the total group to 7 I.Q. points.

The coefficients were:

correlation socio-economic status and 7 year old I.Q.	.374	Significant at 1%
correlation socio-economic status and 11 year old I.Q.	.419	Significant at 1%
correlation socio-economic status and change in I.Q.	.192	Not significant
correlation 7 year old I.Q. and 11 year old I.Q.	.614	Significant at 1%
difference between correlation of socio-economic status and I.Q.s at 7 and 11	.045	Not significant
regression of change in I.Q. on socio-economic status	.141	Not significant

Significance was tested by Fisher's Z transformation.

Neither the regression coefficient nor the difference between the two correlations is significant so the hypothesis is not proved. However, the degrees of freedom possible with differences averaged for schools is dependent on the number of schools not on the number of children. Nisbet (9), also working in Scotland, found that the negative correlation between intelligence test score and family size increased significantly between 7 and 11. Family size is closely linked to social status and it may be that a more extensive investigation or one conducted on an individual rather than a school mean basis would find the regression significant. Extensive rehousing and slum clearance was taking place just before and around the time of the investigation so that a number of schools in the poorest housing areas were closing down. Klineberg (5) has indicated that changes of this kind might have to be taken into account as well as personal factors. A further negative influence was regression toward the mean. The hypothesis required the above average scores to increase and the below average ones to decrease. Chance differences would tend to follow the reverse pattern.

The second hypothesis was tested primarily by the partial correlation coefficients.

The partial correlation coefficients show a significant correlation between socio-economic status and E.Q. with I.Q. constant, but not between socio-economic status and A.Q. with I.Q. constant. The third partial correlation is the most striking. The correlation between socio-economic status and I.Q. with E.Q. held constant is small, negative and not significant. This suggests that the observed correlation between I.Q. and socio-economic status can be attributed to E.Q.

Coefficients were:

correlation socio-economic status and I.Q.	.40	significant at 1%
correlation socio-economic status and E.Q.	.62	significant at 1%
correlation socio-economic status and A.Q.	.52	significant at 1%
correlation I.Q. and E.Q.	.87	significant at 1%
correlation I.Q. and A.Q.	.82	significant at 1%
correlation E.Q. and A. Q.	.90	significant at 1%
partial correlation S.S. and E.Q. with I.Q. constant	.44	significant at 1%
partial correlation S.S. and A.Q. with I.Q. constant	.22	not significant at 5%
partial correlation S.S. and I.Q. with E.Q. constant	—.11	not significant at 5%

E.Q—English Quotient.
A.Q.—Arithmetic Quotient

A factorial analysis was made by Thurstone's controid method. Several iterations were made but were rejected because of the small number of tests. The original communalities were calculated by Medland's Centroid No. 1 (11). Two rotations were made passing the vectors through different tests. Maximum communality was extracted at each rotation and no attempt was made to rotate to simple structure.

Factor Loadings

Unrotated:

	F_{I_A}	F_{II_A}	h^2
I.Q.	.8797	—.1651	.8011
S.S.	.6115	.2426	.4328
E.Q.	.9801	.0515	.9632
A.Q.	.9122	—.1482	.8541
	2.9405	.1107	3.0512

Rotation I:

	F_{I_B}	F_{II_B}	h^2
I.Q.	.8951	—.0003	.8012
S.S.	.5564	.3504	.4324
E.Q.	.9539	.2311	.9633
A.Q.	.9239	.0223	.8541
	2.8743	.1766	3.0510

Rotation II:

	F_{I_C}	F_{II_C}	h^2
I.Q.	.7568	.4779	.8011
S.S.	.6579	.0000	.4328
E.Q.	.9300	.3136	.9632
A.Q.	.7932	.4742	.8540
	2.4996	.5516	3.0511

First and second residuals were tested for significance by McNemar's Test.

Clear cut factors were not expected but it was thought that the factorial model might suggest groupings that mere inspection did not. Two rotations were made each producing two factors. The first rotation showed the grouping of social status and E.Q. These two tests have something in common which the other two do not possess. The second rotation showed a larger common factor and a separate grouping of the three tests—Intelligence, English, and Arithmetic. This seems to suggest that the tests have much in common with social status but something also which is independent. Again English is more closely linked to social status than are the other two tests.

Discussion

The first part of the investigation failed to show a significant correlation between socio-economic status and change in I.Q. between 7 and 11. Other investigators, for example Nisbet (9), have found small changes. Lawrence's (6) investigation showed that the correlation between socio-economic status and I.Q. tends to diminish among children in a residential institution. Apparently a uniform environment tends to iron out previously acquired characteristics. Lawrence also observed that if children were admitted to the institution after their first two years of life the correlation tended to persist. If they were admitted before the second year the correlation was very low. Lawrence's findings are of course, only on the decline of the correlation in abnormal circumstances. Taken with Nisbet's findings they suggest that most of the environmental contribution to the I.Q. is made in the first year though Klineberg's (5) findings, that I.Q.'s of Negro boys who moved from Southern States of America to New York increased, indicate that environmental influences continue to operate.

The second part of the investigation was an attempt to assess the nature of the environmental contribution. The partial correlation coefficients show that when I.Q. is held constant there is significant correlation between E.Q. and socio-economic status but not between A.Q. and socio-economic status. Thus English but not Arithmetic is related to socio-economic status independently of I.Q. That is, of two children of the same I.Q. but different socio-economic status the one of higher social status will tend to be better in English but not in Arithmetic. The third partial correlation shows no significant relationship between I.Q. and social status when E.Q. is held constant. Verbal ability as measured by the English quotient presumably is largely acquired. It is dependent on and limited by

the extent of intelligence but is itself a product of the environment. The correlation then between socio-economic status and intelligence test score can be attributed to verbal factors which presumably are determined by the environment.

Hebb (4) argues that pure intelligence—he calls it intelligence A—is the innate capacity of the brain which shows itself in learned modes of behaviour. The actual intelligence—Intelligence B—used by adults and older children does not involve the formation of new concepts but the use of familiar ones. Even apparently new ideas are built into our patterns of knoweldge in terms of familiar ideas. Consciously or unconsciously we use models for much of our thinking. The physicists' "waves" and "solar systems" in atoms are models that Hebb cites. This second intelligence—Intelligence B—is what our tests test. In this way Hebb provides an explanation of the known cultural factors influencing Intelligence Test scores. Intelligence B apparently is largely dependent on the environment. Intelligence A is not related to the environment but is pure intellectual endowment. If social status is taken as a measure of environmental influences in this group then it is possible to construct a factorial model to fit this hypothesis. On the second rotation all the tests and social status have a large factor loading—Intelligence B; however there is a second factor running through the three tests. They have something independent of the environment—Intelligence A.

The results of the investigation support Neff's (8) thesis that no relationship between socio-economic status and innate potential has been established. There is, however, a real and substantial correlation between socio-economic status and actual intelligence. This split indicates a large wastage of intellectual ability. Children of high potential are held back by stunted development.

The lack of essential stimulation occurs before primary school begins and is not merely a slum phenomenon but occurs throughout the social scale. In this conncetion it is interesting to note that the pupils of three schools were described in almost identical terms by the Headmasters as coming from "solid, interested, working-class homes". In each case I.Q. is high, A.Q. higher but only in one case is the E.Q. as high as the I.Q. These extremely efficient schools have pushed the children's attainment as high as it can be pushed but they have not overcome the basic handicap—a verbal deficit.

A stimulating environment where the cultural deficiencies of the home can be compensated for in part, may be provided in kindergartens. The kindergartens might provide the facilities and opportunities which the underprivileged child lacks. If as Hebb suggests intelligence, Intelligence B, grows in something like

Geometric progression—the more you have the more you can get—then kindergartens are a sound educational investment. Elementary and secondary schools will probably not be able to compensate adequately for the low level of conversation, of entertainment and of cultural interest in many homes but the school curriculum can be planned with more attention to ultimate future achievement and less stress on immediate attainment; so getting the multiple dividend of providing the child not with learning but with the intelligence on which future learning will depend.

Conclusions

1. The correlation between socio-economic status and intelligence has been further confirmed.

2. The correlation between social status and arithmetic can be attributed to the correlation between social status and intelligence.

3. The correlation between social status and English is partly independent of I.Q.

4. The correlation between social status and I.Q. can be attributed to verbal, largely environmental, factors.

5. No significant correlation was found between socio-economic status and change in I.Q. in the primary school.

6. Environmental factors have their greatest effect before primary school begins, therefore attempts to counteract the negative effect of restrictive environments should begin early. These ateempts should have as their main concern the provision of a stimulating environment like that provided by a cultured home with a strong verbal rather than arithmetical emphasis.

REFERENCES

(1) Binet, A. and Simon, T. "Nouvelle recherches sur la mesure du niveau intellectual chez les enfants des ecoles." *Annee Psychol.*, Vol. 17, 1911.

(2) Duff, J. and Thomson, G. H. "Social and Geographical Distribution of Intelligence in Northumberland." *Brit. J. Psychol.*, Pt. II, Vol. 14, 1923.

(3) Fleming, C. M. "Socio-economic Level and Test Performance." *Brit, J. Educ. Psychol.*, Vol. 13, 1943.

(4) Hebb, D. O. *Organization of Behaviour.*

(5) Klineberg, O. *Negro Intelligence and Selective Migration.*

(6) Lawrence, E. M. "An Investigation into the Relationship between Intelligence and Environment." *Brit. J. Psychol. Mon Suppl.*, No. 16, 1931.

(7) Leahy, A. M. "Nature, Nuture and Intelligence." *Genet. Psychol. Monogr.*, Vol. 17, 1935.

(8) Neff, W. S. "Socio-economic Status and Intelligence, a Critical Survey." *Psychol. Bulletin*, Vol. 35, 1938.

(9) Nisbet, J. D. "Intelligence and Family Size", an address to the Conference of the Brit. Psychol. Society, Oxford 1952, reported in the Quarterly Bulletin of the Brit. Psychol. Society, Vol. 3, No. 16, 1952.

(10) Terman, L. and Merrill, M. *Measuring Intelligence.*

(11) Thomson, G. H. *Factorial Analysis of Human Abilities.*

ALBERTA COMPOSITE HIGH SCHOOLS AND GIFTED YOUTH

Ian E. Housego
(Kellogg Fellow, 1957-58)
AND
Gordon L. Mowat
Faculty of Education
University of Alberta

School organization refers, mainly, to the arrangement of facilities and services in an manner designed to achieve an accepted educational objective. One of the lasting objectives in Alberta, as elsewhere ,has been to develop the heterogeneous capacities of children in public schools; therefore, one of the major problems of administration has been to devise a form of school organization—or a system of schools—which will achieve that objective. The most recent organization devised for secondary education is known as the composite high school.

I. The Problem

The purpose of this study was to seek out and evaluate current administrative practices having particular bearing on the education of gifted children in composite high schools in Alberta.

II. Definitions

1. A composite school was defined as being one which offered a complete range of high school courses: matriculation, commercial, technical, and general patterns.

Accordingly this study involved eight composite schools.*

2. Giftedness was defined as being high general intelligence, measured by intelligence tests. For purposes of this study, a gifted student was one who scored in the 97th percentile, or above, on the Dominion Intelligence Test administered by the Department of Education at the Grade 9 level.

It was assumed that the members of the gifted group possessed talents in fields such as fine arts, social relations and mechanics.

III. Method

1. The adequacy of composite schools must be assessed against certain criteria. From literature† dealing with education of the

*Alexandra Composite High School, Medicine Hat; Lethbridge Collegiate Institute; Lindsay Thurber Composite High School, Red Deer; Crescent Heights High School and Western Canada High School, Calgary; Eastglen Composite High School, Strathcona Composite High School, and Victoria Composite High School, Edmonton.

†Including: Brandwein, Cruickshank, DeHaan, Harvard Committee (Report of), Havighurst, Laycock, Martens, N.S.S.E. Publications, Passow, Scheifele and Terman. Also 37 miscellaneous periodicals.

gifted, it was determined that there were generally accepted criteria with regard to the following topics: (1) identification of the gifted; (2) organization within which the gifted receive their education; (3) motivation of the gifted to high achievement; (4) enrichment of the educational experiences of the gifted; (5) co-ordination of school and community resources to facilitate the education of the gifted; (6) supervision of the teachers of the gifted.

2. One hundred and seventeen gifted students were identified from records in the Department of Education.

3. Each school was visited. The investigator studied school records, interviewed all gifted students, guidance officials and principals. Seventy-eight teachers, each having some instructional contact with the gifted student, were requested to complete two questionnaires.

The gross information collected was accepted as being descriptive of present practices in the composite schools.

Section IV, following, develops each criteria, summarizes related practices and presents some evaluations.

IV. Current Practices in Alberta Composite High Schools

1. *Identification*

Theory: Ideally, a secondary school seeks systematically to discover its gifted students. Teachers are supervised in their search; they are provided with objective standards against which they may judge more accurately the capabilities of their students. School authorities administer tests of intelligence, achievement and aptitude. Talents in the fields of music, art, drama, social leadership and mechanics are discovered on the basis of performance in these areas. Performance must be judged by competent persons.

Practice: Alberta composite high schools identify gifted students mainly by means of group tests of intelligence and Department of Education final examinations in Grade IX. Occasionally teachers' judgments are used along with previous school records. Infrequently, staff discussions and standardized achievement tests are used for purposes of identification. While school officials are concerned with the identification of high academic ability, two inadequacies of the total effort to identify are worthy of note:

1. Intelligence testing provided the names of gifted students to the composite high school administrators, but the teachers of the gifted were, on the average, unable to name more than 4 out of 10 of the gifted students.

2. There is very little concern for the discovery of talent in the non-academic areas.

2. *Organization*

Theory: A secondary school striving for maximum development of its gifted students provides an organization which makes possible effective enrichment of their school experiences. Two theories of organization for enrichment are:

1. That enrichment can take place fully only when teachers provide individual attention to the gifted students in "regular" classes.

2. That enrichment will occur most effectively within special ability groupings or by means of the acceleration of the gifted, or a combination of the two.

Current educational theory favours ability grouping and/or acceleration.

Practice: In four of Alberta's composite high schools, conscious efforts were made to group gifted students on the bases of intelligence tests and school achievement records. In all schools there was some degree of ability grouping accomplished indirectly through the course choices of the students. In two schools, administrators opposed the segregation of gifted students from the less able. Several principals favoured acceleration, but viewed certain Department of Education regulations as providing an obstacle to such practices.

Teachers believed that ability grouping would increase opportunities to enrich programmes for the gifted. They contended that present efforts are hampered by heterogeneous classes, by large classes and by lack of time to plan for and teach the gifted.

3. *Motivation*

Theory: A good programme creates a desire in gifted children to achieve high standards of academic work, physical development, and social relationships. Motivation of gifted students may be secured through guidance, extracurricular activities, community contacts and rewards.

The function of guidance goes beyond identifying and making known to teachers, parents, and students the nature of giftedness. Personal guidance is particularly effective with gifted students; it is essential to their maximum development.

The opportunity to participate in extracurricular activities serves to challenge gifted students to develop their talents. A good programme provides a wide range of "tryout" opportunities such as drama clubs, school orchestras, science clubs, school newspapers and athletic teams.

Gifted students who are brought into close contact with gifted adults in the community may be inspired to emulate these adults in matters of academic excellence, performance and aspirations.

Awards for high achievement and outstanding contributions to school life will motivate gifted students to a remarkably high degree. Such awards may take the form of honour rolls, membership in honour societies, prizes, and scholarships.

Practice: The composite high schools used the results of intelligence tests, vocational guidance tests, and to a lesser degree, aptitude tests, in gaining a knowledge of the "gifts" of their talented students. Personality tests and standardized achievement tests were used very rarely. All the schools made reference to the cumulative record folders of the gifted students, each of which contained considerable information on the history of the student's previous schooling.

Gifted students visited guidance counsellors for interviews approximately twice each year. The interviews centered around educational and vocational guidance. Very little guidance of a "personal" nature was provided. According to the students, the influence of guidance counsellors on their course choices and extracurricular participation was very limited.

Ordinarily, guidance counsellors informed students of their potentialities and the students were urged to work to capacity. Under-achieving gifted students received special guidance in only a few schools.

Looking at the whole extracurricular activities picture, the eight schools provided opportunities for the development of interests and abilities in 165 clubs, associations and atheltic activities. In addition, each school had a Students' Council and issued school publications of the newspaper and yearbook types. These provided positions of responsibility for willing and able students. The average number of extracurricular activities offered per school was 21. The most commonly offered activities were athletic in nature. Relatively few opportunities appeared to be found in the academic and fine arts areas.

Only two schools out of eight consistently provided gifted students chances to meet appropriate, talented individuals in their respective communities. In four schools, members of the community met the gifted students, along with the whole student body, during "Career Day", when it was hoped some vocational orientation might be achieved.

In seven out of eight schools annual awards were made to students of high academic achievement in Grade X, XI, and XII. In

all schools there were awards for students who had made outstanding contributions through extracurricular activities. There were available national, provincial, and local scholarships. Five of the eight schools provide recognition in the form of monetary rewards.

The attitudes of gifted students towards school may reflect the net motivation being achieved.

1. Sixty-four percent of the selected students had a "fairly strong" liking for school; seventeen percent had a "very strong" liking for school; seventeen percent had only a "slight liking"; two percent "positively disliked" school.

2. The most-liked subjects were in the mathematics-science area; those most disliked were in the languages-social studies-literature area.

3. Ten percent of the selected students felt overworked in the composite high schools; fifty percent said they were underworked. On the average, the selected students spent approximately 10 hours a week on school studies outside of school hours.

4. Approximately one student in ten experienced failure during his high school career. Predominantly, failures occurred in Grade XI and in one subject only.

5. The majority of the selected group participated in extracurricular activities, predominantly of an athletic nature.

6. Less than half the group won awards for outstanding achievement in either academic or extracurricular activities.

7. Approximately eighty percent of the group planned to carry on with further education at universities, colleges, and other places of higher learning. Engineering, teaching, and nursing were the most frequent vocational choices.

4. *Enrichment*

Theory: A curriculum designed for gifted students offers experiences more intensive and more extensive than are suitable for average students. It is reasonable to expect them successfully to meet the challenge of school work which demands high ability. Therefore, courses are provided within a wide range of electives, especially designed to challenge and develop a variety of giftedness. A curriculum of this nature requires that gifted students make extensive use of library and laboratory facilities.

Teachers who enrich their courses should at least plan and organize activities that involve the following procedures: small group and individual research; seminar discussions on problems of special interest; more difficult and demanding assignments than the average

are required to do; widespread reading of a relatively difficult nature and the reporting of some reading to class.

Administrators must provide the teacher of the gifted with small classes, with planning time, and with supervisory assistance.

Practice: The data gathered suggested that the teachers and administrators, as a group, recognized the special curriculum needs of gifted students. Principals were willing to (and did) facilitate enrichment in and out of the classrooms; still, there was no policy or programme of enrichment formulated specifically in terms of the needs of gifted students. The majority of teachers stated that they provided some degree of enrichment in their classes. They appeared to be "on their own" in this respest, and received little or no direction from administrative authorities. A few of the teachers of the selected group did outstanding jobs in challenging gifted students. Most teachers, if they provided enrichment at all, provided very little.

The great majority of gifted students were enrolled in the matriculation programme. Some of the schools organized what might be termed "honour classes" in English, French, and mathematics, which were restricted to academically-able students.

Research indicates that gifted stuents read more widely and read more difficult books than their lessable classmates. This study showed that approximately one-quarter of the selected students never made use of composite high school libraries.* The remainder did use the libraries while doing assignments and for general reading enjoyment. Nearly half the students felt that the library facilities were less than adequate in meeting their reading needs; a similar proportion of teachers agreed.

What was generally true of the libraries also was true of laboratory facilities. A very limited number of the gifted took part in individual research and small-group projects.

The credit system in Alberta opens a way for a degree of enrichment by allowing students to carry "extra" courses. The average number of credits earned by gifted students during their high school career was 108 as compared to a maximum of 120 possible credits. One hundred credits are required for a high school diploma.

One must conclude that composite schools are not enriching adequately the curriculum of the gifted.

5. *Co-ordination of School and Community Resources*

Theory: There are two ways in which secondary schools concerned about the education of gifted youth make use of community resources. Firstly, they encourage the gifted to take part in worth-

*This does not include their possible use of home and public libraries.

while community organizations, because high schools are not always in a position to provide the type of enrichment desired. Choirs, libraries, art classes, and countless youth organizations are particularly effective community outlets. Secondly, the schools may exploit the community, through field trips to museums, courts, concerts, commercial firms and industrial plants. Conversely, interesting citizens may be invited into the school to meet gifted students, to talk and to work with them.

Practice: The data showed that gifted students, as a group, were relatively active in one or more community organizations, but not evidently under guidance and direction from the school. The organizations to which they belonged were mainly of service or athletic nature.

Nineteen percent of the group did not belong to any community organizations. Presumably these students had interests, aptitudes and abilities that community organizations were in a position to develop. The schools were making no planned effort to encourage students to take advantage of opportunities available.

6. *Supervision*

Theory: Secondary school administrators concerned with the provision of proper education for gifted students ensure adequate supervisory leadership of the teachers of the gifted. These teachers are involved in a programme of in-service education, specifically designed to help them understand the characteristics of gifted students, their peculiar needs and problems, and to acquaint the teachers with up-to-date ideas and practices concerning the education of gifted students.

The supervisory programme is concerned also with building an enriched curriculum for the gifted. All the teachers are involved in planning and the execution of the curriculum and understand, therefore, its aims.

Furthermore, to be effective, the supervisory programme must include evaluation of teachers and facilities. This requires constant examination of a school's total approach to the education of its gifted students.

Practice: In-service training of teachers as "teachers of the gifted" was generally non-existent. Some of the teachers were involved in curriculum revision at a provincial level; some were involved at a local level. In either case, none was involved in curriculum building directed specifically toward the education of the gifted.

The majority of teachers stated that they were practicing some curriculum revision in the classroom in an attempt to partially meet

the needs of their gifted students. What they did seemed to rest on their initiative rather than upon a planned, school-wide effort.

Composite high school authorities had no policy of evaluating teachers of the gifted as such; at least, none of the teachers was aware of having been so evaluated.

No more than one-third of the teachers questioned had attempted to increase their knowledge concerning the education of gifted students.

V. Conclusion

The criteria used in this study imply an ideal which may not be achievable wholly in a single given school. Therefore, evaluation in terms of this ideal may fail to estimate accurately, relative to the potential of a particular school, the degree of effort being devoted to education of the gifted. Certainly, attitudes of the community, the school board and the staff may impose limitations upon the total effort of any school. It is equally certain that a measure of provincial attitude is reflected in present practices in composite schools.

Nevertheless, if one accepts the assumptions upon which this study proceeded, the evidence is somewhat disconcerting. The study did not seek to establish whether there was uniformity of approach to educating gifted children; rather, it sought evidence which might indicate clarity of purpose.

The latter is the focal point of one conclusion: the composite schools, as major centers of secondary education with relatively favourable resources in facilities, curriculum and personnel, have a largely latent responsibility for the education of gifted children.

REFERENCE

(1) Housego, Ian Edward "Alberta Composite High Schools and Gifted Youth", unpublished thesis, University of Alberta, 1958.

AN INTRODUCTORY STUDY OF THE CONSISTENCY OF MARKING MATHEMATICS PAPERS IN SCHOOLS

D. H. Crawford
Faculty of Education
University of Alberta

Introduction

This article is a summary of a research study carried out by the author some years ago while teaching in Scotland.

The author knows of no subsequent study in this field as far as Mathematics is concerned, although research on the marking of English Essays continues apace, at least in England. An excellent summary of such research up to 1956 is given by Vernon et al. (11)

The study was initiated with a view to obtaining general data on the consistency of marking secondary school mathematics papers, and, in particular, in order to obtain answers to the following two questions:

(1) Do the marks awarded by different markers to the same set of papers vary significantly from one marker to another?

(2) Does the consistency with which the same marker marks and remarks the same set of papers vary significantly from one marker to another?

The study was so designed as to indicate further possible research on marking. In view of the current interest in "The Meaning of Marks", (6) the publication of this article may prove helpful in initiating further studies.

Survey of Previous Research

A survey of research work relevant to the investigation was made. The studies considered can be conveniently grouped in the two categories below.

Mathematics Marking

The work carried out on the marking of 23 University Mathematical Honours Papers by Hartog and Rhodes (5) was reviewed. The most noteworthy points of their study were these.

1. The extreme difference between the mean marks awarded by any two of the six individual examiners was 3.7%. Improvement to a 3% level of difference was obtained when markers marked in pairs.
2. Disagreement as to the merit of a given candidate ranged from 5% to 21% for individual marking. For pair marking, the range was 1% to 15%.

3. The mean level of the above disagreement (i.e., differences between extreme marks) was 11.6% for individual marking and 6.1% for pair marking.

This investigation gave some idea of the degree of variability of marking mathematics papers obtained in practice, and what reduction in variability is obtained by having papers marked by a team of two examiners.

English Essay Marking W

The studies in this field which proved most helpful were those of Thomson and Bailes (9), Hartog (4), Hartog and Rhodes (5), Cast (2), Wiseman (12) and Finlayson (3). A consideration of these threw light on the meaning of the term consistency, and revealed the main statistical techniques employed in recent years to investigate it.

Consistency. Two types of marker consistency were revealed. The first, named *inter-consistency,* is essentially a comparison between different markers relative to some fixed standard which is external to each individual marker, e.g., the correlation coefficient of the marks of each marker with the pool of the rest of the markers. The second type, named *self-consistency,* is measured by the correlation coefficient for each marker between his two sets of marks when he marks and then re-marks the same papers. A third type of consistency—pupil consistency from one test to another was also revealed by a study of Finlayson's work.

Statistical Techniques. The main statistical techniques employed in previous research were found to be

(1) the calculation of means and standard deviations
(2) the calculation of correlation coefficients
(3) the analysis of variance
(4) the factor analysis of correlation coefficients

The most comprehensive research on essay marking, that of Finlayson (3), demonstrated how to find the overall reliability of any examination, taking into account all the variables concerned, including the varialibilty of the children and of the examination itself. Although this would have been the ideal experimental design for the author's study, various reasons made it impossible to carry it out.

Experimental Procedure

The raw material for the investigation consisted of the summer examination scripts of 35 first year pupils of a Scottish Senior Secondary school (mean age 13 years, 2 months; standard deviation 6 months). There were eight questions in the examination paper;

three on Geometry, three on Algebra, and two on Arithmetic. Of a total of 100 marks, 35 were allotted to Geometry, 38 to Algebra, and 27 to Arithmetic. 210 pupils from six class sections sat the examination, and a representative sample of 35 scripts was then selected from this total. The mean and standard deviations of the marks of the population and sample were 53.73 and 53.43, and 16.8 and 19.0 respectively. All signs of the normal marking for school purposes (which had been done in pencil) were removed as well as the pupil's name, and the scripts were then identified by the allotment to each of a mumber from 1 to 35. Each of the scripts was then marked by each of 8 markers on two occasions, with an interval of three months between marking and re-marking. Of the 8 markers, 4 taught in the same school as the pupils attended (but the maximum number of scripts drawn from any class taught by any of the markers was 6). None of the remaining 4 knew each other, or any of the first 4. Each group of 4 consisted of two men and two women, and each member of the second group was chosen to be approximately equivalent in experience to a member of the first group. Each marker was provided with (1) a standard marking sheet subdivided into columns, so that question totals, subject totals, and examination totals could all be analysed if desired, and (2) a sheet of marking instructions, which ensured that each of the 16 markings was uniform, and uninfluenced by any previous marking.

Analysis of the Marking

A. *Analysis of raw marks, means and deviations.*

The average marks awarded by each marker overall, and in each subject separately, were calculated, and from these Table I was drawn up.

(1) It will be seen that the average levels of marking remain very consistent over the two markings with a maximum difference of 0.8% in the case of Geometry. (2) The average of the deviations of the individual marker means from the general mean is of the order of 2-2.5% overall, and is least in Geometry and greatest in Algebra.

A table was also compiled to show the complete list of marks awarded to each pupil by each marker at both markings. The marks of pupil number 25 ranged from 44% to 66% in the first marking, and from 41% to 69% in the second (same markers in each case). The spread of marks of 25 of the 35 pupils was 10% or more in both markings, while in the case of 7 pupils, (i.e., 20% of the sample) disagreement existed between certain markers as to whether the pupil should pass. These findings underline the necessity of taking care in marking border line cases.

TABLE I

DIFFERENCES IN STANDARDS OF MARKING

Subject	Marking	Mean Mark of 8 Markers in %	Average % Deviation from the Mean	Maximum % Difference between Means of any two Markers
Overall	1	50.7	2.4	9.7
	2	50.9	2.1	9.5
Geometry	1	51.4	1.9	9.2
	2	52.2	1.8	6.0
Algebra	1	52.3	3.8	19.1
	2	52.5	2.8	14.6
Arithmetic	1	47.5	2.1	10.4
	2	47.1	2.5	9.6

B. *Correlation Methods.*

A few product moment correlation coefficients between different markers were calculated. They were all nearly unity (.98 or .99); a very different result from the values found in essay marking (of the order of 0.6). This shows clearly that the markers agree very closely as regards the order of merit of the candidates, but it fails

TABLE II

MARKERS VS PUPILS ANALYSIS USING 35 PUPILS

Source	d.f.	Sum of Square	Mean Square (Variance)
Markers (M)	7	2187	312.5
Pupils (P)	34	157531	4633.3
Error (M × P)	238	2521	10.6
TOTAL	279	162239	

to account for the difference in general level of marking, and some other statistic which took account of this might prove valuable.

C. *Analysis of Variance.*

Due to lack of time, this analysis was applied to the first marking only.

The fractions of the total variance due to the markers, to the pupils, and to residual error, were calculated. The analysis involved is presented above in Table II. The F ratio of marker's variance/error variance (29.6) then gave a criterion of whether the markers varied significantly from one to another in their marking. We are in fact testing the null hypothesis that the marks awarded by the markers differ by chance only. Since a value of $F = 2.72$ is sufficent to refute this hypothesis by demonstrating differences which would not arise in more than 1% of cases, we can confidently say that the differences in marking are extremely unlikely to be due to chance alone.

After random exclusion of two of the 35 pupils, the remaining 33 pupils were divided into three equal groups of 11, representing high achievement, medium achievement, and low achievement on the basis of the total of the 8 marks awarded to each by the eight markers. Analysis of variance was then carried out on each group separately as before. High significant F ratios (10.2, 20.5, and 12.3, where $F = 2.91$ for a probability of 1%) were again obtained. Thus it appears that the markers differ significantly in their marking not only overall, but with pupils of different levels of achievement. The fact that the F ratio is largest for the marking of the group of average pupils suggests that a closer study of marker consistency (or variability) in this region would be worthwhile.

Thirdly, a similar analysis was carried out for the marks awarded in Geometry, Algebra and Arithmetic separately. Once again highly significant F ratios (7.5, 32.9, and 11.9 where $F = 2.72$ for 1%) were obtained showing that the significant variation in marking persists in the individual subjects, with the suggestion that it is greatest in Algebra.

Several further analyses were carried out, of which *two* may be mentioned briefly. The four markers from the same school were treated as one group, and compared with the other group of markers to see if any significant difference in marking were present. The F ratios obtained were negligible, thus justifying the underlying assumption of the enquiry that the markers had been chosen at random.

The last analysis to be mentioned here was a grouping of the markers into two groups according to sex. Variances between sexes

and within sexes were calculated. An F ratio of 1.49 (d.f. = 1/278) was obtained for the group of 35, showing that overall there were no significant sex differences in the marking. However, a larger study of this problem of possible sex difference in marking might prove worthwhile.

Analysis of Self-Consistency

Early investigators used as a measure of self-consistency the mark-remark correlation of a marker's marks in two markings of the same scripts. As already noted, agreement on order of merit in the study under consideration was about perfect, which would yield a conclusion that each marker is perfectly self consistent. An attempt was therefore made to investigate self-consistency by another method, namely, using a modification of Bartlett's Test of Homogeneity of Variance (1). Briefly, if x_{ij} and y_{ij} denote the marks awarded by marker i to pupil j on the first and second markings respectively then $Si^2 = \frac{1}{k-1} \sum_{j=1}^{k} (x_{ij} - y_{ij})^2$ where k is the number of pupils, is a measure of the variance or lack of self-consistency of marker i.

The method then consisted of a comparison of k times the logarithm of the average variance with the sum of the logarithms of the separate intra-marker variances by means of the formula $\chi^2 = (k-1) \left[n \log_e \bar{S}^2 - \sum_{i=1}^{k} \log_e S_i^2 \right]$. Reference to tables of χ^2 with the appropriate degree of freedom then gave a criterion of whether the markers were equally inconsistent in their marking or not.

As before, the analysis was carried out for the 35 pupils as a group, then for the three achievement groups and finally for the three separate subjects.

The results of the first and third analyses were all significant at the 1% level showing that the degree of inconsistency exhibited by a marker over the two markings varies significantly from one marker to the next. Break up of the pupils into three achievement groups yielded values of χ^2 which were either not significant or barely so. Why this should be so was not clear, and in view of the small numbers in each group, not much weight should be given to this finding, without further research.

Note. An alternative method of investigating self-consistency using analysis of variance was subsequently suggested to the author.

This analysis was not carried out but is given here for reference as it may prove useful in similar investigations. It consists essentially of combining the marks awarded by a given marker in the two markings, using these new marks to obtain the sum of squares between markers, finding the total sum of squares of the ungrouped marks, and then, by subtraction, finding the sum of squares within markers. Variances for between markers and within markers can then be calculated and an F ratio obtained by treating the within markers variance as residual error. If the value of F obtained is significant, this shows that the variance between markers is significantly greater than that within markers and "intra-variability" may be neglected in comparison with "inter-variability".

Graphicial Representation of Consistency

The total mark awarded by each marker (a) in each subject (b) to each achievement group, was taken, these eight marks were summed, and their mean and standard deviation found. Each total mark was then expressed as a standard score (i.e., as a deviation from its particular mean in units of standard deviation). Graphs were then drawn for each of the two markings and are shown in Figures 1 to 4. These graphs give a general picture of both types of consistency. They show, for example, whether a given marker occupies the same position relative to the mean when marking different groups or subjects. Any idiosyncrasies of individual markers are thus shown up. Thus marker H is generally consistent, and the level of her marking is close to the mean (Figures 1 and 2). She is most variable in marking Geometry, and this is true of marker F also. The graphs appear to show that markers mark Arithmetic and Algebra similarly (the lines for these stay together) but not Geometry. Marker C shows the characteristic of marking high achievement pupils more leniently than he does other groups (Figures 3 and 4).

This method of presentation has the advantage of showing many results at a glance, and might prove very useful in picking out particular characteristics required in a marker.

Summary

The main results of the study may be summarized as follows:

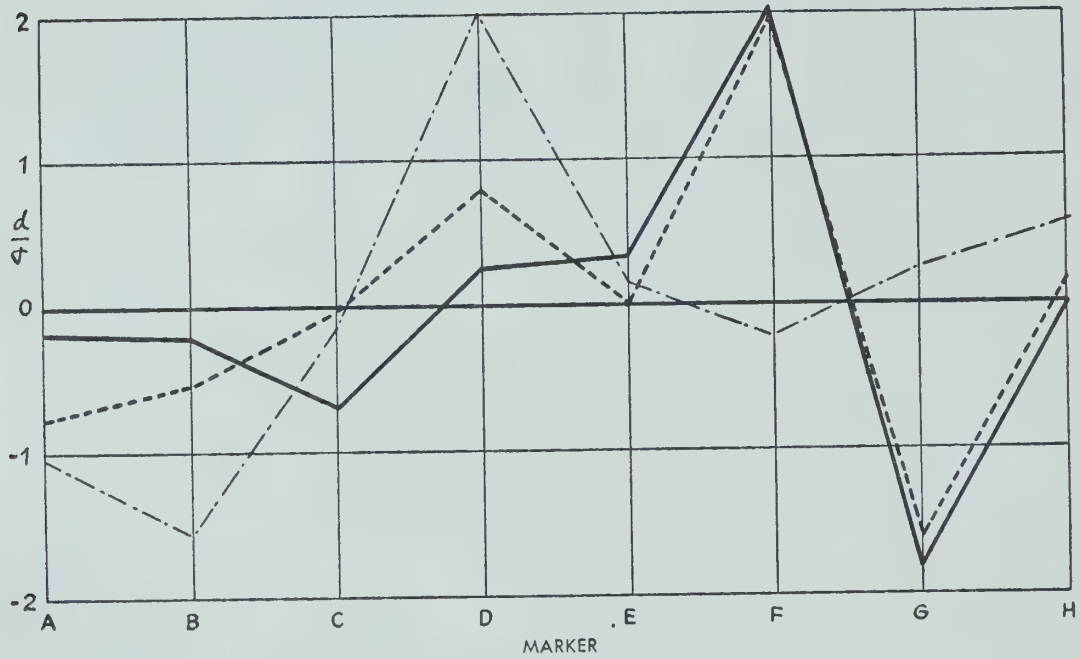
1. Different markers in mathematics award marks which differ in a manner greater than that to be expected on the basis of chance alone, individual pupils being awarded marks often differing by 10% or more. This suggests that for school and other examinations, it is important to develop a system of marking which is fair to each student, e.g. Marker A marks *all* answers to questions 1, 3, and 5,

MARKS AWARDED IN DIFFERENT SUBJECTS
(IN UNITS OF S.D. ABOUT MEAN)

GEOMETRY ALGEBRA ARITHMETIC

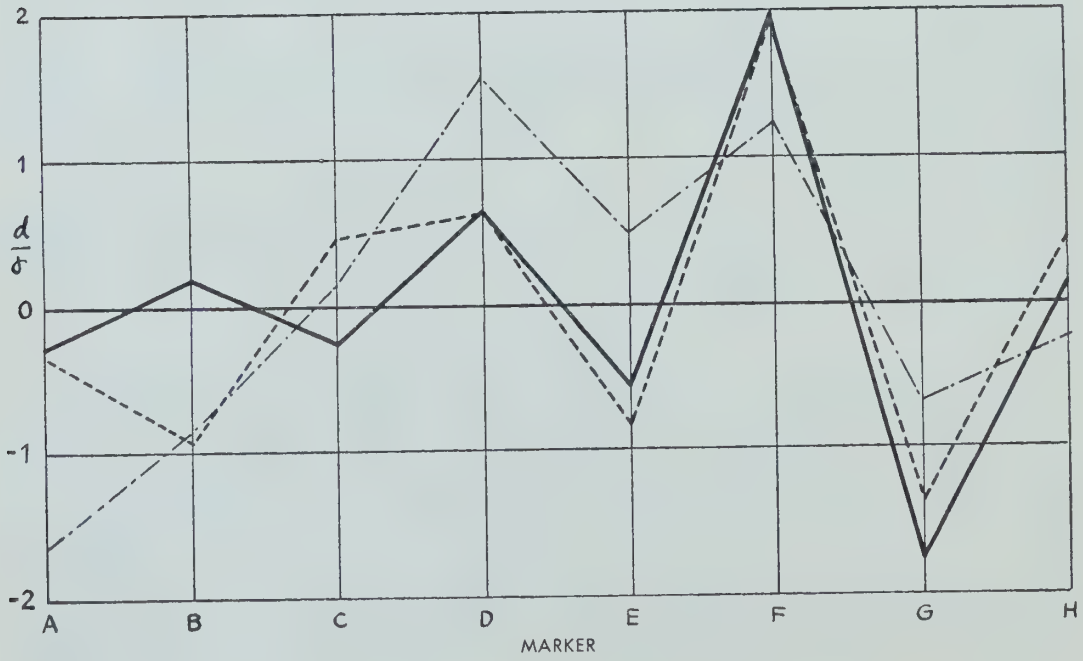
MARKING 1

FIG. 1



MARKING 2

FIG. 2

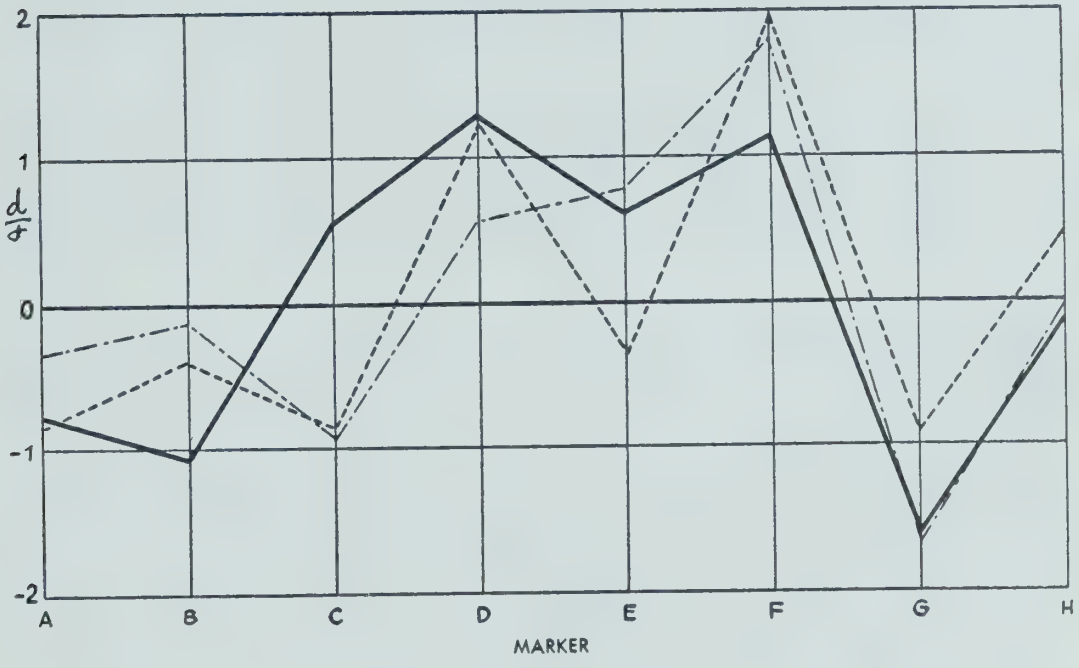


MARKS AWARDED IN MARKING GROUPS OF DIFFERENT ACHIEVEMENTS
(IN UNITS OF S.D. ABOUT MEAN)

HIGH ACHIEVEMENT — MEDIUM ACHIEVEMENT - - - LOW ACHIEVEMENT

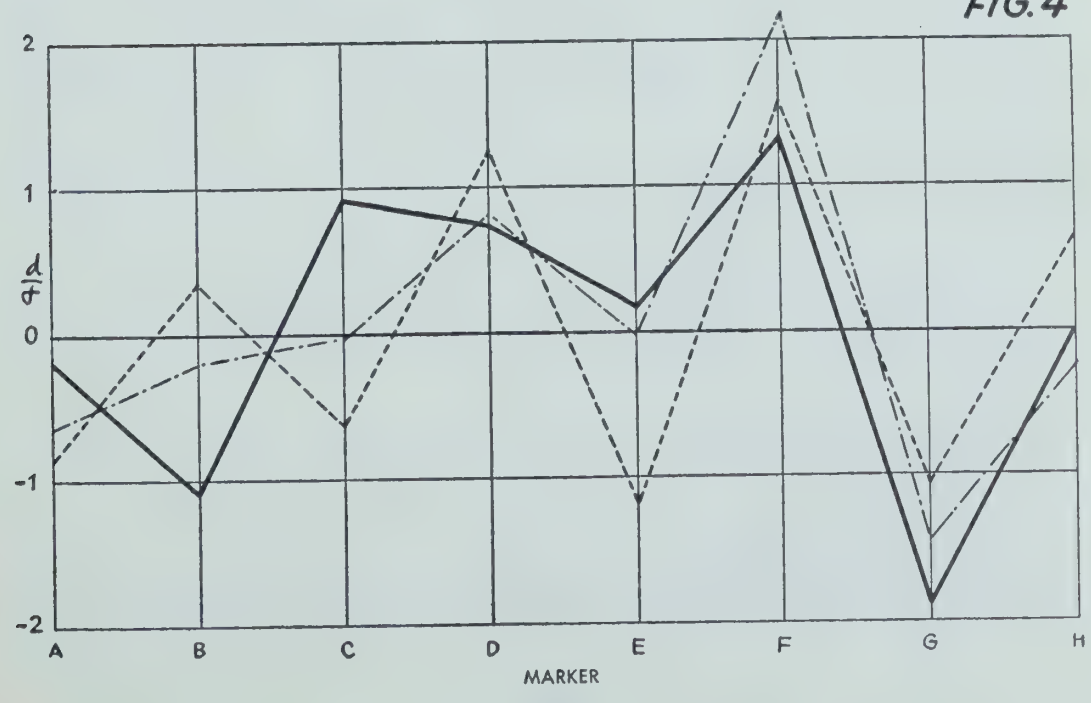
MARKING 1

FIG. 3



MARKING 2

FIG. 4



and Marker B *all* answers to questions 2, 4, and 6 *OR* Marker A marks *all* Geometry questions, Marker B *all* Algebra questions.

2. The degree of self-consistency exhibited by markers was found to differ significantly from one marker to another. It is clear that a marker with a high self-consistency is a more reliable marker, and this appears relevant to the selection of markers to mark examinations in which standards of difficulty and passing are meant to remain constant from year to year. Further research is indicated.

3. The significant variations in both types of consistency persist when Geometry, Algebra and Arithmetic are marked separately, the variation appearing to be greatest in the case of Algebra. Some fairly rigid scheme of marking seems necessary to reduce these variations .

4. Similar significant variations are found when the marking of groups, of low, medium and high achievement is considered. Further study of marking the "border-line" achievement group is particular is indicated.

5. Graphical anlysis of consistency shows advantages in presentation of results, and can prove helpful in showing up individual marker characteristics and trends in marking.

REFERENCES

(1) Bartlett, M. S. *Supplement to the Journal of the Royal Statistical Society,* 4, 1937, P. 137.

(2) Cast, B. D. M. "Efficiency of Different Methods of Marking English Compositions." *British Journal of Educational Psychology,* 7, 1939, 257-269, and 8, 1940, 49-60.

(3) Finlayson, D. S. "The Reliability of the Marking of Essays." *British Journal of Educational Psychology,* 21, 1951, 126-134.

(4) Hartog, P. et al. *The Marking of English Essays.* London: Macmillan, 1941.

(5) Hartog, P. and Rhodes, E. C. *The Marks of Examiners.* London: Macmillan, 1936.

(6) Hocking, B. "On the Meaning of Marks." *University of Alberta Research Note* (unpublished), 1958.

(7) Linquist, E. F. *Statistical Analysis in Educational Research.* Boston, New York: Houghton Mifflin Company, 1940, Ch. 4, 5 and 6.

(8) Snedecor, G. W. *Statistical Methods.* Ames, Iowa: Iowa State College Press, 1946.

(9) Thomson, G. and Bailes, S. "The Reliability of Essay Marks." *Forum of Education,* IV, 1926, 85-91.

(10) Thurstone, L. L. *Multiple Factorial Analysis.* Chicago. University of Chicago Press, 1947.

(11) Vernon, P. E. (ed.) *Secondary School Selection.* London; Methuen and Company Limited, 1957.

(12) Wiseman, S. The Marking of English Composition in Grammar School Selection. *British Journal of Educational Psychology,* 25, 1951, P. 200ff.

ACHIEVEMENT AND UNDER-ACHIEVEMENT IN ENGLISH LANGUAGE TEN IN AN ALBERTA COMPOSITE HIGH SCHOOL

ROBERT STANLEY CHAPMAN
Calgary, Alberta

The Problem

Recent changes in the Alberta curriculum have included more teaching time for English language. The traditional language courses, which were keyed to formal analysis, have been replaced by those which stress the functional use of language in social situations. All of this increased change and stress points toward an awareness in Alberta of the importance of language as *the* core subject.

Despite this interest on the part of teachers and departmental officials, there are many students of both low and high intelligence who are not achieving up to their capacities in language. What factors differentiate between the under-achiever and the student who is making the most of his abilities? How can these factors be isolated: (a) as a means of predicting achievement, and (b) with a view toward removing some of the causes of under-achievement? This study throws some light on these questions, with particular reference to one grade-level in one large Alberta high school (2).

Experimental Procedure

For each of the 471 Grade X and XI students registered in English Language 10 in Strathcona Composite High School, Edmonton, on May 1, 1957, an achievement ratio was found by dividing an index of performance in English by an index of mental capacity. Only 448 Grade X students constituted the final sample. The achievers were defined as those 112 students who had the highest achievement ratio, and the under-achievers as those 112 students who had the lowest achievement ratio.

The null hypotheses that there were no differences between the two groups were tested with respect to the following dimensions:

(1) Personal data: (a) sex, (b) intelligence, (c) chronological age.

(2) Previous school achievement: (a) Grade IX language mark, (b) acceleration or retardation.

(3) Interests: (a) pattern of Grade X courses, (b) study habits, (c) participation in extra-curricular activities, (d) out of school clubs, (e) hobbies, (f) private lessons, (g) after school jobs.

(4) Background data: (a) Language spoken at home, (b)

parents' occupation, (c) junior high attended in Grade IX, (d) parents' schooling.

(5) Personality Ratings: (a) self-ratings, (b) teacher ratings.

In order to find his achievement ratio it was necessary to find for each students—(i) his capacity index, and (ii) his performance index.

(i). Capacity index. The capacity index used was Mental Age as computed from the General Test administered on a province-wide basis as one of the compulsory Grade IX final examinations of June, 1956. This was the Dominion Group Test of Learning Capacity—Intermediate—1950 Edition—Form A—Alberta Edition. The General Test raw scores for 431 Grade X students in attendance May 1, 1957 at Strathcona Composite High School, were listed from the Department of Education Examination Branch records. Mental ages were computed by the method described in the "Guide to the Interpretation of the Grade IX General Test—Administered June 1956". Twenty Grade X students (through illness, a move into the province, etc.) had not written this General Test. As part of the school guidance program, seventeen of the students in October, 1956 had completed the Otis Group Test of Mental Ability (Form X) and the computed M.A. from this test was used. Three of the students who had not written the General Test, also did not write the Otis Test, and were not included in the sample.

The seventeen Grade XI student taking Language 10 fell into two groups (1) those who were repeating because they had failed it or wished to raise a low mark and (2) those who had attended a very small high school in Grade X where Language 10 and 20 were taught in a two year cycle, and thus had not had an opportunity to take Language 10 previously. The M.A. as computed from the General Test in their Grade IX year was used for all Grade XI students. These Grade XI students were not included in the sample but were discussed separately.

The remaining 448 students constituted the final sample.

(ii). Performance index. In light of the previous research done in this field it was felt that the measure of performance should be composed of measures of both mechanics of English language and effectiveness of expression of English language. Standardized tests, teacher-devised course content tests and samples of written expression were used.

Mechanics of language was measured by the Co-operative English Test A—Mechanics of Expression—Form X, and by the Strathcona Composite High School Spring Language 10 examination. The latter test, a multiple-choice objective type, was devised jointly by

the six Language 10 teachers of Strathcona Composite High School, to test the coverage of the course content from September 4, 1956 to May 1, 1957. In a split-halves study, a reliability coefficient of .67 was found for this test.

Effectiveness of expression was measured by the Co-operative English Test B 1—Effectiveness of Expression—Form X, and by a teacher-marked essay written at the same time as the Spring Language 10 examination.

Objective test scores and the essay scores were transmuted into T-scores, (with a mean of 50 and a standard deviation of 10). By means of the tables provided by the publishers, the Co-operative tests were changed into scaled scores, (with a mean of 50 and a standard deviation of 10). To give equal weighting to all four measures, these four standardized scores were added to give the individual student's performance index.

This performance index was then divided by the capacity index (mental age expressed in months) to give his achievement ratio. *It should be noted that this achievement ratio was not a measure of intellectual capacity, but rather a ratio of performance in English to mental capacity for each student.* Some students of high intelligence fell below the lowest quartile of achievement ratio, and conversely, some students of low intelligence fell above the highest quartile of achievement ratio.

The students who fell above the upper quartile of the resultant achievement ratio scale were classed as those who were achieving near or up to their capacities and are called "the achievers" in the remainder of this study.

The students who fell below the lowest quartile of the achievement ratio scale were not achieving a high as their capacity measure suggested they might and these people are referred to in the remainder of the study as "the under-achievers".

In addition to writing the four previously mentioned tests (used to determine achievement ratio), all 468 Language 10 students completed (1) a Brown and Holtzman Survey of Study Habits and Attitudes, and (2) a questionnaire (including a personality and temperament self-rating scale) devised by the investigator.

Teachers were asked to complete a similar rating scale for selected students who fell into either the achiever or under-achiever category. The students were not identified to the teacher-raters as achievers or under-achievers.

The rating scale was adapted from that used by Beaton (1) who in turn used the personality rating scale devised by Terman (3) in his later studies of the development of gifted children.

Summary of Findings

1. *Personal data*

The achiever group was composed of 82 girls and 30 boys with a mean chronological age of 16 year 0 months and a mean mental age of 15 years 11 months. The under-achiever group was composed of 21 girls and 91 boys with a mean chronological age of 16 years 3 months and a mean mental age of 15 years 7 months.

2. *Previous School Achievement*

(a) While 69% of the achievers received a mark of H or A on the Grade IX departmental language examination, only 37% of the under-achievers received a mark of H or A on the same examination.

(b) Only 7 of the achievers had ever failed while 38 under-achievers had failed at least one grade. Ten achievers and 4 under-achievers had been accelerated.

3. *Interests*

(a) Of the achievers, 71% were registered in the academic pattern of courses and none was registered in the shop pattern. Of the under-achievers, 53% were registered in the academic pattern and 29% in the general pattern.

(b) The achievers had significantly better study habits and attitudes than the under-achievers on the Brown-Holtzman Survey of Study Habits and Attitudes Scale. Table I presents the distributions of the two groups on this measure, and provides an example of distributions for which the significance of differences in means and variances was later computed.

(c) Athletics was by far the most popular extra-curricular activity for all students. Twenty per cent of the achiever girls took part in the modeling club. The majority of the students spent one hour per week on their activity. Twenty-two achievers and 6 under-achievers held an executive position.

(d) Religious and teen clubs were the most popular out-of-school organizations, to the achievers. Athletics, teen, religious and cadet organizations in that order, were joined most often by the under-achievers. The achievers spent two to three hours and the under-achievers three or more hours on these activities. Thirty achievers held office (quite often a major position) and 25 under-achievers held office (usually as leaders of small sub-groups).

(e) While only 28% of the achievers spent over three hours per week on their hobby, 48% of the under-achievers devoted that much time. The male under-achievers deviated from the usual hobbies

of only sports or some sedentary activity and were interested in cars, airplanes, radio and electricity.

(f) Most of the 34 achievers who took private lessons spent from 2 to 8 hours: 15 under-achievers took such lessons, devoting the same amount of time to them.

TABLE I

DISTRIBUTION OF RAW SCORES ON BROWN AND HOLTZMAN SURVEY OF STUDY HABITS AND ATTITUDES

Score	Achievers	Under-Achievers
55-59*	1	1
50-54	4	1
45-49	6	1
40-44	5	6
35-39	16	7
30-34	20	12
25-29	14	12
20-24	16	27
15-19	15	19
10-14	14	17
5- 9	1	8
0- 4	0	1
Total	112	112
Mean	28.12	22.80
S.D.	11.30	10.55

*The higher scores indicate better study habits.

(g) Exactly 50% of the girl achievers and 50% of the boy achievers had jobs: 12% spent ten or more hours at their jobs. A total of 46% of the girl under-achievers and 57% of the boy under-achievers worked after school (26% of the boys spent ten or more hours). There was an overwhelming affirmative response to the question, "If you do not have an after school job now, would you like one if you could get one?"

4. *Background Data*

(a) While 13 achievers came from foreign-speaking homes, only 4 came from homes where the foreign language was spoken most of the time. Twenty-one under-achievers came from such homes (including 9 Ukrainian and 7 German). Eight signified that the language was spoken most of the time.

(b) More under-achieving boys (28) came from the "best" district than from any other single district. Twenty-seven of the achievers came from the 'second best' district.

(c) Fifty per cent of the achievers had a parent who was a professional person or the owner or manager of a business. Twenty-one per cent of the under-achievers' parents came from this grouping. Twenty-seven per cent of the under-achievers came from the 'labor' group.

Thirty-four per cent of the mothers of girl achievers were employed and 43% of the mothers of girl under-achievers were wage earners. In the same category, the boys were: achievers—13%, under-achievers—24%.

(d) The median amount of education for mothers and fathers of both groups was "some high school". There were 42 fathers with more education than this (including 11 university graduates) whose children were achievers, compared to 30 fathers with more education than "some high school" (including 5 university graduates) whose children were under-achievers. The number of mothers of achievers with educational qualifications above "some high school" was 45 (including 7 university graduates). Mothers of under-achievers with education above "some high school" numbered 31 (including 3 university graduates). Table II presents the distributions of the two groups on this measure, and provides an example of distributions for which the significance of difference was later computed using the X^2 test of independence.

TABLE II

EDUCATION OF PARENTS OF ACHIEVERS AND UNDER-ACHIEVERS

Education	Achievers		Under-Achievers	
	Father	Mother	Father	Mother
Did not attend school	2	0	8	7
Public school	32	29	33	32
Some high school	32	34	38	35
Graduated from high school	15	15	9	21
Some training beyond high school	16	23	16	7
University degree	11	7	5	3
Total	108	108	109	105

5. *Personality Ratings*

Both achievers and under-achievers rated their own personalities as very stable, and about average in all aspects except friendliness, in which they rated themselves slightly above average. The ratings

of achievers by teachers agreed with the students' self-ratings but the ratings of under-achievers by the teachers did not agree with the self-ratings. The teachers rated the under-achievers as "rather quiet", rather sensitive", "slow to adopt to change", and "inclined to give up too easily".

Conclusions

The various null hypotheses of no differences between the achievers and the under-achievers were tested using the T-test for differences in percentages, the T-test for differences in means, the F-test for differences in means, and the X^2 test of independence; the 0.01 level was adopted for significance in each instance. When a significant difference in means was found, the distributions concerned were tested for homogenity of variance before decision was made as to the significance of the difference in means. Table III summarizes the conclusions reached with regard to the statistical significance of the differences between achievers and under-achievers.

The achievers, in comparison with the under-achievers, were more often female, were younger chronologically, had higher marks in Grade IX language, had failed less often between Grades I and IX, were taking an academic course, had better study habits and attitudes, spent less time on out-of-school hobbies, took more private lessons, came from a professional or semi-professional family, had parents who had more formal education, and were rated by themselves and a teacher as having a stable personality.

The under-achievers were more often male, and were older chronologically. They had lower marks in Grade IX language, a possible failure between Grades I and IX, and poorer study habits. The under-achievers were taking an academic, a general or a shop course (in that order), and spent more time on out-of-school hobbies. They took few private lessons, came from a non-professional family, and had parents who had less formal education. The under-achievers felt that they had stable personalities but the teacher-raters disagreed in some aspects and described them as rather quiet, rather sensitive, slow to adapt to change and inclined to give up too easily.

No significant differences were found between the two groups in the following dimensions: mental age, amount of acceleration, amount of time spent on extra-curricular activities, amount of time spent on out-of-school clubs, amount of time spent on after-school jobs, language spoken at home, or Junior High School attended in Grade IX.

No such person as the "average" achiever or under-achiever

TABLE III

STATISTICAL SIGNIFICANCES OF DIFFERENCE BETWEEN ACHIEVERS AND UNDER-ACHIEVERS

Dimension	Statistical Test Used	Conclusions, in direction of Achievers
1. *Personal Data*		
(a) Sex	t. Diff. of percentages	Sig. More girls
(b) Chronological age	t. Diff. of means	Sig. Younger
(c) Mental age	t. Diff. of means	Not Sig.
2. *Previous School Achievement*		
(a) Grade IX Language	t. Diff. of percentage	Sig. More H's or A's
(b) Retardation	t. Diff. of percentages	Sig. Fewer previous failures
(c) Acceleration	Very few cases, not tested	
3. *Interests*		
(a) Pattern of Grade X courses	t. Diff. of percentage	Sig. More in academic pattern
(b) Study habits	t. Diff. of means	Sig. Better
(c) Time on extra-curricular activities	F. Diff. of means	Not Sig.
(d) Time on out-of-school clubs	F. Diff. of means	Not Sig.
(e) Time on hobbies	F. Diff. of means	Sig. Less time
(f) Private lessons	t. Diff. of percentage	Sig. Took more private lessons
(g) Time on after-school jobs	F. Diff. of means	Not Sig.
4. *Background Data*		
(a) Language spoke at home	t. Diff. of percentage	Not Sig.
(b) Junior High School attend	X^2. Independence	Not Sig.
(c) Parents' occupation	X^2. Independence	Sig. More often professional or owner of business
(d) Parents' education	X^2. Independence	Sig. More years of formal schooling
5. *Personality Ratings*	(Not tested statistically)	
(a) Self-ratings	Modes and Profiles	Teacher ratings agreed with self-ratings for Achievers, disagreed with self-ratings for Under-Achievers.
(b) Teacher ratings	Modes and Profiles	

probably exists, and the investigator does not suggest that *individual* students be labelled, for example, Girl—therefore probably an achiever, good study habits, etc., or Boy—therefore probably an under-achiever, poor study habits, etc.; but the differences indicated might be useful in predicting the achievement of *groups* for purposes of special groupings, special curricula and remedial instruction.

Many of the factors which seem associated with under-achievement cannot be changed, but others may provide opportunity to raise the under-achiever from that category. These include pattern of courses (some under-achievers are out of their depth and not in the field of their interests); study habits (many are building a mental block of dislike, and others do not know how to study); and time spent on out of class activities (the ten or more hours spent by 26% of the male under-achievers probably detracted from out-of-class school work).

The fact that the two groups were spread over the range of intelligence from below average to superior supports the view that performance in school work should be reported not only in relation to the rest of the group but also in relation to the individual's capacity, i.e., a student may be performing in the middle or even top portion of the class but may be performing quite unsatisfactorily in relation to his capacity. Conversely, a student in the lowest portion of the class may be working to the utmost and be doing very well in relation to his capacity. The implications for the hopes of parents and the vocational plans of students are obvious. The achievement ratio found by dividing a suitable performance index by a suitable capacity index can be an important measure.

REFERENCES

(1) Beaton, M. A. "The Effects of Acceleration on the Academic Progress and on the Personal and Social Development of Calgary and Edmonton Grade 10 Students." Unpublished Master's thesis, University of Alberta, 1954.

(2) Chapman, R. S. "A Study of Achievement and Under-Achievement in English Language Ten in an Alberta Composite High School." Unpublished Master's thesis, University of Alberta, 1958.

(3) Terman, L. M. *Genetic Studies of Genius.* Stanford University Press, 1926.

TEACHER EDUCATION IN ENGLAND

H. S. Baker
Faculty of Education
University of Alberta

The road from the monitorial system, that "first desperate approach" to teacher education, to the modern Area Training Organization has been a long one. For the most part it has been travelled slowly. But more rapid movements in the twentieth century have, as usual with English education, followed closely upon major Committee reports.

The McNair Report, (4) published in 1944, dealt at varying length with most matters pertaining to the supply, status, and training of teachers. Under the latter head it made positive recommendations about organizations, institutions, and programs.

The Committee found that "existing arrangements for the recognition, the training and the supply of teachers are chaotic and ill-adjusted even to present needs." (4, p. 18) Though the training colleges, 83 in all,* were theoretically in liaison groups under the Joint Examination Boards of ten universities, there was little co-operation among them. Twenty-nine were operated by local education authorities. Fifty-four belonged to voluntary bodies—Church of England, Roman Catholic, Methodist, and non-demoninational groups. Sixty were for women, sixteen for men; seven were co-educational. Most of the colleges were uneconomically small, poorly equipped, and inadequately staffed.

Physically at any rate the 22 university training departments were thought to be better off. With shorter traditions, however, they suffered from lack of prestige within their universities, and like the training colleges were lacking in intercommunication.

Area Training Organizations

The McNair Committee were in clear agreement on the first step toward reform. This lay in the appointment of a Central Training Council of from three to five members which should "assume the obligation, not of taking over the existing training colleges or of establishing new ones, but of ensuring that training institutions adequate in number and quality, are available and are fused into a national training system." (4, p. 45) The Council's initial duty was to be a survey of the needs and resources of each area with a view to recommending groupings "on educational grounds."

*But not including sixteen schools for the training of art teachers. There are now more than 150 training colleges—some 50 of these being run by voluntary organizations, and 100 by local education authorities.

Committee opinion was divided, however, as to the place of the university in the training organization. Exactly one-half of its ten members were convinced that the university, "because of its recognized standards and its standing in the educational world . . . must be the focus of the education and training of teachers in the future." (4, pp. 50, 52) They therefore recommended:

> (a) that each university should establish a School of Education, it being understood that some universities may find it desirable to establish more than one school;
>
> (b) that each University School of Education should consist of an organic federation of approved training institutions working in co-operation with other approved educational institutions; and
>
> (c) that University Schools of Education should be responsible for the training and the assessment of the work of all students who are seeking to be recognized by the Board of Education as qualified teachers. (4, p. 54)

The other half of the Committee (including the chairman), being of opinion that the main duty of universities "consists and properly consists in teaching basic subjects and in the advancement of knowledge," (4, p. 55) urged that more fruitful relationships would result from continued use and extension of the Joint Boards. "The University training departments and the training colleges should preserve their separate identity." (4, p. 57) Again, "we are often told that our schools are suffering from the too academic outlook of many of their teachers, and we are opposed to a policy which would tend to emphasize that outlook." (4, p. 61 and note) The second group of five therefore recommended, in part:

> (a) that the Joint Board . . . reconstituted should become responsible for the organization of an area training service in which there will be a university training department and training colleges preserving their identity and being in direct relation with the Board of Education and the Central Training Council . . .
>
> (f) that the Joint Board should be responsible for the examination and assessment of students both in the university training department and in the training colleges . . . (4, p. 62)

In the actual working out of the above recommendations, organizations following the former became known as Type A, and those following the latter as Type C. Initially there were only three of the latter: Reading, Liverpool, and Cambridge. Now there is only one, Reading and Liverpool having become Type A.

Previously the national Board (now the Ministry) of Education had exercised rather close control over teacher training, scrutinizing with some care the work of the Joint Boards before approving it. Clearly, then, the acceptance of the area training idea (whether Type A or Type C) involves the yielding of many prerogatives and powers. But the Ministry still controls staffing, over-all intake, grants, buildings and equipping.

Institutes and Institutions

Initially the term *institute* was referable to Type C training organizations. Subsequently, however, it came to be used indiscriminately with reference to both types. In at least one instance it has been further used to designate a single institution—the Central Training College of the London Institute of Education.

The McNair Committee urged a vastly extended scope for the training organizations.

> In the first place, institutions other than purely training institutions, such as technical colleges, colleges of art, colleges of music and agricultural institutes, should be brought within the ambit of the training service and contribute to the education of students preparing to be teachers. In the second place, admission to training should not be limited to those who either at the end of a secondary school course proceed to a training college or at the end of a degree course seek a one-year course of training. All forms of post-secondary education, such as that provided in technical colleges, colleges of art, agricultural institutes, adult education classes and such as that obtained independently at home or elsewhere, should be regarded, under reasonable safeguards, as a proper qualification for a course of training. In the third place, courses of training should not be limited to a three-year course for those admitted without a degree and a one-year course for graduates. The training service should provide courses, suitable in nature and length, to meet the needs of applicants showing great variety as regards age, maturity, attainments and experience. (4, p. 63)

The major institutions of the area training organization remain, of course, the *training colleges* so-called, and the *university departments of education.* Unlike the university departments, the training colleges are regularly provided by the local education authority. Even those voluntary colleges which have retained their independence require a measure of supervision and approval by an area training organization before their students can be recognized as qualified teachers by the Ministry of Education.

> The Two Year training Colleges are perhaps the most characteristic of our arrangements for the training of teachers, and the oldest of them constitute the original nucleus of the whole system. Their function, which has changed remarkably little since the earlier of them were established in the thirties and forties in the nineteenth century, is to provide a combined course of general education and professional training leading up to an examination upon which the student may qualify for recognition as a Certificated Teachers. (3, p. 17)

Characteristically, the training colleges now enrol non-graduates for the two-year program for initial qualification, though many of them offer also a one-year program for mature students,* and a few of them offer a one-year program for graduates. Occasionally there are to be found third-year or supplementary courses in special

*A development of the Emergency Training Program set up to meet the teacher shortage in the years following the Second World War, and quite in accordance with the principle of flexibility recommended by the McNair Committee.

subjects (science, domestic subjects, commercial subjects, etc.), and even a three-year course for combined degree and certification.

The university's role in teacher education is in general more recent. At the present time the major enrolment of the university departments is that of university graduates for one year for initial qualification—the certification program. There are also diploma programs of one year for experienced teachers (psychology, divinity, the deaf, youth service, etc.), as well as those leading to the Master's and Doctor's degrees.

Selection is an important feature both of the student's choice of college and the college's choice of student. The latter submits his application, together with a list of institutions in order of preference, to the college of his first choice. If he is not accepted here, the application is forwarded to colleges of the second, third, or subsequent choices until he is accepted by one or until the list is used up. In the latter event the application goes to "the pool," a clearing house administered by the professional body of the training college lecturers. If a suitable vacancy occurs the student may yet be successful.

For their part, most colleges and departments select their students with very great care. The student must, of coures, first meet the academic requirements—for the university department, a degree; for the training college, stipulated G.C.E. requirements.* He is then likely to be called for an interview with the first principal, either at the college or at or near his home.

Training and Qualification

As might be expected, the training colleges slant most of their work toward the lower levels of the school program: Nursery, Infant, or Junior.** Most of them do, however, offer Secondary Modern programs, and a few offer Secondary Grammar.

The main concentration in the university departments is on Secondary Modern and Grammar School programs, with the frequent addition of Secondary Technical and Commercial and also of Junior School programs.† A few university departments provide for the Infant School as well.

Though the area training organizations do not attempt to dictate the structure and details of the training program inter-communica-

*Five subjects at ordinary level, four at ordinary level and one other at advanced, two at ordinary and two other at advanced, or three at advanced level. These requirements may be relaxed on the authority of the Area Training Organization.

**Individual, or variously grouped—as for example Nursery/Infant, Infant/Junior, Junior/Secondary.

†Cambridge, a Type C organization, offers programs only for Grammar and Public Schools. Oxford caters for Preparatory, Secondary Modern, Secondary Grammar, and Public Schools.

tion and a good measure of consensus do in fact lead to substantial similarities. The courses are said to be "professional, general, and semi-professional," and are to be "kept in appropriate relation to one another." (15, Pt. II, Ch. V)

The professional group always includes a course on the principles of education. In the training colleges this is likely to be broad rather than deep, with attention to the philosophy and history of ideas, psychology, sociology, and administration. In the university departments, on the other hand, philosophical, psychological and historical foundations are almost always handled as separate courses; whereas in the training colleges the extra emphasis tends to be on English and physical (health) education. The latter often appears as a specific course in university departments.

Methods courses, termed "semi-professional" because the subject matter involved it itself regarded as liberalizing, tend to be organized both specifically and generally. In a training college, for example, the student might choose the teaching of arithmetic *or* religious education *plus* that of a group such as literature, drama and social studies. The breadth of the group varies with the level: at the Infant level the group is likely to be very broad, while at the Secondary level a single (second) subject may be substituted. In a university department there is often a "general methods" course in addition to the special choice, the latter being specific to a school level (Secondary, Modern, Junior, etc.) or to a subject. The most frequent among the latter are English, history, geography, mathematics, physical science, biological science, music, modern languages, classics, religion.

The "general" subjects are so called because they are regarded as relative to the student's general education as well as, or in deed rather than, contributory to his fund of information for teaching. Popular offerings in the training colleges, where the student elects one or two, are as follows: geography, history, mathematics, music, biology, physical education, arts and crafts. Since in the university departments the student is a university graduate he is not normally required to extend his general education. But he may, in consultation with his tutor, add one or two "options" of either general or professional significance—for example, speech training, visual aids, drama, music, history of educational ideas, intelligence testing, development of science, physical education, arts and crafts.

Teaching practice, is of course, an important part of the training program. One of the "drastic" proposals of the McNair Committee had been the extension of this phase of the work to one continuous term. This was, however, contingent on the addition of a third year to the training college program. At the present time most training

colleges and university departments provide 110 half-days or the equivalent, grouped into periods of observation and various kinds of teaching practice under the guidance of training staffs and qualified teachers. It is usual to give each student some experience at a school level other than that at which he aims to specialize.

Qualifying procedures vary somewhat from institute to institute, and even from institution to institution within the same area. It can be said that, with the insistence of the local education authorities on practical classroom competence, there has been some shift away from the more limited academic assessments on the Joint Boards. At the moment recognition is given for achievement in ongoing course work as well as in final examinations, and certainly for practice as well as for theory. Examiners are internal as well as external to the student's institution, and even at times to the institute. The assessment of the external examiners is final.

Attention should be called to the fact that a university graduate is automatically qualified. It is significant that in 1957 only 53 per cent of new teachers were trained—as compared with 80 per cent over-all.

Some Current Issues

Student Life

In comparison with the graduate students in the university departments of education, training college students live a secluded life. This is hardly surprising in view of the religious, even monastic origins of many training colleges.* Small size, scattered location, residential setting, single sex, limited program—all these features tend to perpetuate their seclusion.

The tradition of segregation has gone for the most part unquestioned until recent years. Even the 1926 Committee was inclined to be defensive of professional and sex groupings.

> During professional training we see no reason why teacher-students should not be trained in institutions where other students, if any, are a minority. (3, p. 86)
>
> Apart from the probable administrative difficulties connected with the accommodation and staffing, there will normally be complications arising from differences in the curriculum for men and women, especially in small colleges . . . (3, p. 102)

The McNair Committee in 1944, however, looked guardedly toward change.

> We have more than once referred to the segregation of students in training colleges by profession and, in most cases, by sex. Our recommendations would go a long way to break down segregation college by college; and in course of time it may be that the area training services

*Monopolized by the churches until 1870.

> generally will provide, as some individual colleges now do, for the education and training of those who do not intend to be teachers. This process cannot be hurried, but any proposal, in response to an effective demand, to provide courses for students preparing for another profession should be encouraged. Our proposals will also help to break down segregation by sex. Our witnesses were not unanimous on the question of "mixed" training colleges, but we have no hesitation in saying that the proportion of institutions based on the principle of co-education should be increased. (4, p. 75)

Undoubtedly the area training organizations are having something of the effects anticipated, though even now the ratio of men's to women's to coeducational colleges is in the order of 1:10:5½. Whether this situation—in view of the fact that most teachers-in-training themselves come from single-sex secondary schools, and that more and more of them will teach in mixed schools—can be longer defended is surely questionable and questioned.*

A Third Year for Training Colleges

> The ordinary Two Year course necessarily covers a wide field, and there can be little or no time for a student with a special bent or aptitude to cultivate it, as it ought to be cultivated, in the interests of future pupils. . . . The aim (of a third year) will be to increase the teacher's professional competency, either by means of enlarged knowledge of the subject or by a study of the best current methods of teaching it, or by both (3, p. 100)

The general argument thus formulated in 1926 has been rather steadily pressed ever since. The McNair Committee in 1914 complained that "many students in the training colleges do not mature by living: they survive by hurrying." (4, p. 65) In 1956 the National Advisory Council on the Training and Supply of Teachers urged that "the institution of a three-year course of general training is the next step in the development of the highly qualified teaching profession which is required to meet modern educational needs." (9, p. 3) In 1957 a Ministry pamphlet argued that the extra time should be "used boldly for particular purposes." (12, p. 1)

What needs and purposes? A recent statement by the London Institute of Education indicates the ways in which "teachers would be more fully equipped for service in the schools:

> (1) they would have a better training in the Theory and Practice of Education and for that reason be better able to appreciate and serve the educational purpose of the schools;
>
> (2) they would have a more fundamental training in method, thereby making them more adaptable to the variety of methods used in the schools;
>
> (3) the general teacher would be well qualified to teach in an wider range of subjects;
>
> (4) there would be fuller opportunities for the abler students to pursue their special interests and to qualify themselves to teach subjects at a higher level, especially if, as seems probable, Colleges are able to exercise a greater measure of selection in the admission of students;

*University Departments of Education are, of course, regularly coeducational.

(5) students would be able, especially in the third year, to achieve a greater measure of maturity as persons and as teachers. (16, p. 12)

The foregoing and similar statements would commend themselves to most educators—with reservations, perhaps, on the part of a few who regard all professional education as superfluous to general education. Indeed, the third year as such is hardly an "issue" on merit. But what may make it so are emerging conditions of teacher supply and demand, and of national finance. As to the first, it had been estimated that conditions should make it possible to introduce the third year in or about 1961 without prejudice to staffing demands or to the continued improvement of teacher-student ratios. (9, pp. 7-11) There is now some question, however, as to whether these hopes have been too sanguine. There is also the fear that new national policies of financial retrenchment may delay or prevent the extra year.

General and Professional Education

What has been called "the doctrine of personal development" is deep-rooted in English teacher education. Not only is it felt that teachers become good teachers by virture of a good general education and of a good specific education in the subject or subjects they will teach, but it is further felt that great value accrues from the intimate, personal study of a subject or subjects "in depth", even though this may be quite unrelated to the teaching field.

The doctrine has not gone unchallenged, or at least unqestioned. In 1926, for example, the Departmental Committee wrote as follows:

> The reform introduced in 1913 appears . . . to have brought with it some disadvantages . . . The choice allowed in academic subjects seems to have been almost too encouraging in some instances; the possibility of carrying the study of chosen subjects to a comparatively high point appears sometimes to have absorbed an almost too generous share of energy . . . They (Training College students) are interested in children rather than subjects. In their course the study of children being the central idea, subjects should be studied from the point of view of the part they play in a child's education and of how they can be presented. Such study is educational in itself and not a mere acquisition of skill . . . We desire to see the course more professional, but not less educative. (3, p. 91-93 *passim*)

This questioning is recurrent.

> Is a dichotomy between the educative and the professional function of the college a valid distinction? . . . We may say that the professional needs will be more adequately met because the teacher is a more educated person because of his study. What also needs to be said is that the more specifically professional courses can be concerned in the same way as instruments of the student's general education . . . (14, p. 8)

The foregoing statements refer to the training colleges. If the argument for professionalism is valid there, it should logically be

more so in the university departments, to which students come with degrees in general education and in special subject fields. Thus:

> The year must be a humanizing one . . . (but) it can do much to open the mind of the student to sound and possible methods of teaching which otherwise he might never encounter . . . What we want is to get a man's mind exercised as an educating mind, not only as an academic mind . . . The important questions of method are all questions of principle. (13, p. 10)

Theory and Practice

The concluding sentence (above) points on toward a related issue: to what extent is professionalism in the sense of classroom competence the result of theory, and to what extent is it the result of practice? Both training staffs and students have tended to urge the primary importance of teaching practice, and the allocation of more training time thereto. Despite this fact it has been argued that illusions of competence often result from mere classroom habituation, and that such illusions may blur the need for a disciplined study of principles. Or to put the matter in another way:

> The undue emphasis in some training colleges on "teaching practice" to the neglect of problems of theory, and the difficulties of establishing education as a university discipline, are symptomatic of . . . traditional views. (6, p. 5)

The University and Teacher Education

The Departmental Committee in 1926 had high expectations of the contribution that the universities could make to the work of teacher training.

> Proposals directed to emphasizing the professional efficiency of the Training Colleges do not preclude bringing them into closer relationship with the universities. The universities stand for the values of thinking and knowing, of intellect and knowledge, in the national life. It is natural that any institutions concerned with these values should look to the universities for standards and inspiration, and we cannot doubt that they on their side will always be ready to co-operate. (3, p. 103)

The cooperation hoped for was of two main kinds: some university credit for the training college program (with even some participation in lectures), and university advice and assistance in the councils of training colleges and other educational bodies. The first hope has not been fulfilled. The second, however, has been substantially met.

With reference to the latter it is of course necessary to ask about the nature and direction of the university influence on teacher education. All Institutes except Cambridge are now Type A. This kind of relationship is thought by some to be a "potential source of danger." (15, p. 112)

> If they (the universities) attempt to make conditions—in a field which is in many ways different from that of the university proper—conform too

> rigidly to the pattern of university routine, tradition and convention, and if too great a weight is given to the opinion of those who are insufficiently informed, the training institutions may find that they have acquired a closer relationship with the universities at too great a price in terms of their freedom. But if the universities use their power discreetly, and with a full realization of the repository of knowledge and experience which exists in the training institutions and in the teaching profession, then the future is indeed hopeful. (15, pp. 122-123)

Superficially at least the above fears seem groundless. University and training college staffs appear to enjoy generally cordial relationships. This could mean that the two groups do in fact meet in a spirit of mutual education and (where necessary) compromise. It could, however, mean any one of a number of other things: for example, that the training college staff (or, less likely, the education department staffs) have become submissive junior partners in the enterprise of teacher education, or that the education departments themselves are submissive to the universities at large.*

With increasing years, one way of assessing the value of the Institutes will be the measure of their interest in and impact on something much broader than the mere production of secondary school teachers—public education generally, "the scientific investigation of all those influences that tend to modify the behavior of the human individual and to shape the course of his development . . ." (7, p. 180)

> This kind of study of education can be properly pursued only in a university with the best resources of knowledge and personnel that the university can supply. It has a significance far beyond the teaching profession, but it is the essential basis of teacher training. That is the fundamental reason why the universities could not have evaded the responsibility they have now accepted. (7, p. 180)

What the above writer does not say it that all teacher education should actually go forward within a university, the education department assuming supervision of the total program (both general and professional). The concept, however, is neither new or forgotten.**

Some Implications for Canadian Education

Selection

A notable characteristic of English and Continental education is the high degree of selection at all levels beyond the elementary. It is a matter of "pupils" and "places", with the latter almost invariably fewer than the former. This is in rather sharp contrast to North

*Each Institute is governed by a council consisting of representatives of the university, the training colleges, local education authorities, and the teachers associations. Academic matters, however, are mainly or completely under university control.

**W. H. G. Armytage, for example (1), sees its development as an outgrowth of area training arrangements, as follows: phase 1 (1929-44)—Joint Board System; phase 2 (1944-57)—Area Training Organizations; phase 3 (195?-?)—a Bachelor of Education degree of "pass" or "general"type for all teachers. The University of North Staffordshire operates a program of conjoint professional and liberal arts courses during the last three years of a four-year undergraduate degree, similar in many respects to that of universities in the United States and Canada.

American education, in which (with the exception of a few preferred schools and faculties—Medicine, for example) institutions tend to take all comers who can meet minimum entrance requirements, and to deal with them as best they can.

Canadian high schools have more or less successfully met the "bulge" under this liberal dispensation. It remains to be seen whether the universities can or will do the same. (One school of thought would select rather than expand, or at least select more and expand less.)

In considering the alternatives it seems important to distinguish between general and professional education. If it is true that the more broadly educated an electorate the better, then selection should have a minimum place in general public education. But it by no means follows from this that professional education should be similarly open. One may reasonably argue that if teaching is or should be a profession, and that if this profession is of great public value—as indeed all profess it to be—those permitted to prepare for and to practice it should be a select group.

Canadian educators may therefore look somewhat enviously toward England, where many institutions for teacher education do in fact select their students. That is to say, they do in fact interview, assess, choose and reject. However, the disillusioning corollary is that on a national scale the familiar North American pattern tends to reappear—with its well known characteristics of supply, demand, and minimum requirements. Ultimately some less preferred institution (be it the third, fourth, or lower choice) will "select" students of minimum qualifications.

True selection can operate only when society is prepared to place a value on education and to underwrite that value in financial and status terms. This much is obvious from Canadian experience alone. No expedient will do.

A more difficult, though not less important aspect of selection is that of character, personality, and suitability for the more intimate requirements of the teaching job. To what extent and for what particular reason should candidates as persons be accepted or rejected? The English practice seems rather more effective, since interview procedures do at least direct attention to qualifications that transcend the merely academic. Again, however, there is no assurance that candidates personally unsuited to teaching may not at the last moment be dredged from the pool.

Extent

Here teacher education in Canada clearly suffers by comparison with that in England. Few of us would defend the English practice of extending qualified-teacher status to university graduates with

no professional study, or to assert that the one-year program for graduates is superior to any of the Canadian patterns of preparation for teaching in secondary schools. In England, however, the standard two-year program for elementary teachers is a general minimum. It is true that many Canadian teachers also follow a two-year basic program, and that some provinces encourage a four-year degree program for all. But the persistence of one-year and other "emergency" programs in some provinces would be intolerable in England.

If Canadian school standards are in any way inferior to those of England, here is surely one place to look for the remedy.

Teacher-Student Ratio

And here is another. The actual, not theoretical ratio of instructor to students in teacher-training institutions in England is one to eleven. In Canada the fraction would be distinctly less favourable. The significance of this difference—with all that it means for the guidance of student teaching as well as more general kinds of guidance and instruction—is not to be minimized.

The University and Teacher Education

Provincial Departments of Education in Canada retain a much tighter control of teacher education (as of curriculum and most other aspects of public education) than does the English Ministry of Education. In approximately one-half of the Canadian provinces the Department itself operates the training colleges for elementary school teachers, approving the university—through its school of education—for the training of secondary school teachers. In rather less than half, the university Faculty has been appointed for the training of all teachers. In each of the latter instances, however, the university operates under clear if indirect pressure from the provincial Department, whose authority at the policy level must be termed major. It is further significant that the province retains control of certification.

What of English policy and practice? Assuming for the moment that the Canadian provinces were prepared to yield control to Institutes of Type A, what would be the effect upon Canadian teacher education?

The difficulty of knowing the success of the Institutes—in terms of cooperation, purpose, general effectiveness—has already been suggested. Perhaps it is too early to assess them. They have certainly improved intercommunication (between the university as such, its department of education, the training colleges, the local authority, etc.), and they may have stimulated profitable controversy

in education. Similar results have been achieved, however, by Canadian Faculties of Education.

It is with reference to educational controversy that the university's role becomes critical. Productive action demands a measure of community, and in the event that differences are severe it may demand an arbiter—or a leader. Is the university best qualified, best disposed to play this role? If so, should it be the university as an institution, or its faculty of education? Theoretically, traditions of free thought and scholarship suggest the university's qualification for the task. Yet the objectivity or impartiality of university thinking has not always extended to matters of public education; indeed it is not so long ago that North American universities, at any rate, abdicted from teacher education as beneath their academic dignities.

Much of the problem obviously turns on the old and difficult question of political control. It is clear that the English state assumes a much lighter prerogative in education than does the Canadian or the American. In a very real sense the heavy prerogative in England lies with educational leaders and the teaching group. It has been suggested, and as vigorously denied, that the same situation might obtain on this Continent.

Dare the schools build a new social order, as with Counts? Dare the universities—directly or otherwise—control the teaching order and, through this, the social order? Is it an acceptable proposition that any politically unresponsible body should legislate in matters so crucial as those of public education? Is there, indeed, a future for Dewey's idea of the school state in which neither governments nor educators but rather the educative process itself shall be instrumental in determining social directions and values?

REFERENCES

(1) Armytage, W. H. G. "Training Colleges and Universities in the Third Phase." *The Universities Review* 29:2 (February 1957), pp. 56-59.

(2) Association of Teachers in Training Colleges and Departments of Education. *Handbook on Training for Teaching.* London, Methuen, 1954.

(3) Board of Education. Report of the Committee on the *Training of Teachers for Public Elementary Schools.* London, H. M. Stationery Office, 1926.

(4) Board of Education. Report of the Committee on *Teachers and Youth Leaders.* London, H. M. Stationery Office, 1944.

(5) Hockey, Stanley. "The University Professional Training Course Today." *Journal of the Institute of Education of Durham University* 6 (March 1955), pp. 12-14.

(6) Holmes, Brian. "The Teacher of Teachers." *Education for Teaching* 32 (May 1954), pp. 3-14.

(7) Jeffery, G. B. "The Universities and the Teaching Profession." *Universities Quarterly* 1:2 (February 1947), pp. 178-184.

(8) Ministry of Education. *Challenge and Response.* Pamphlet No. 17. London, H. M. Stationery Office, 1950

(9) Ministry of Education. *Three Year Training for Teachers.* Fifth Report

of the National Advisory Council on the Training and Supply of Teachers. London, H. M. Stationery Office, 1956.

(10) Ministry of Education. *Training Colleges in England and Wales.* London, H. M. Stationery Office, 1957.

(11) Ministry of Education. *Training and Supply of Teachers.* First Report of the National Advisory Council on the Training and Supply of Teachers. London, H. M. Stationery Office, 1951.

(12) Ministry of Education. *The Training of Teachers.* Pamphlet No. 34. London, H. M. Stationery Office, 1957.

(13) Nibblett, W. R. "Teacher Training and the Graduate Today." *Journal of Education* 86:1014 (January 1954), pp. 10-12.

(14) Tibble, J. W. "The Training Colleges and the Three-Year Course." *British Journal of Educational Studies* VI:I (November 1957), pp 3-12.

(15) Unesco. *The Education of Teachers in England, France and U.S.A.* Paris, Unesco, 1953.

(16) University of London Institute of Education. *The Scope and Content of a Three-Year Course of Teacher Training.* London, the Institute, January 1956.

(17) *Year Book of Education.* London, Evans Brothers, 1953.

EDUCATIONAL RESEARCH

Journal Review

Volume One, Number One, of a new research journal has just come to hand: *Educational Research,* a publication of the National Foundation for Educational Research in England and Wales. The appearance of a new journal reporting educational research is an interesting but not a spectacular occurrence. This journal is sufficiently different from existing journals to warrant regarding its publication as something of a special occasion.

Educational Research will not report on specific research projects. Rather it is intended to provide 'comprehensive and authoritative statements of the principal facts and conclusions derived from research,' and to make them available in intelligible form for teacher and laymen. The first editorial recognizes the difficulties which beset an individual who tries to keep up with educational research published in the research journals. The sheer volume of published work is more than any busy teacher can read and digest. In addition refinements in statistical and experimental methods essential to the soundness of the work done inevitably increase the difficulty of understanding for those who are not themselves professional research workers.

This journal is designed to bridge the gap between research workers and teachers. The articles in each number will be 'substantial summaries of work in various research fields.' Topics of broad importance and massive content will be treated in serial fashion, reviewing the research over a number of issues of the journal. Emphasis will be placed on practical problems of teaching method and organization. Single articles on general problems will vary the offering.

In addition to the articles the review will have two continuing features which are certain to prove useful. In each issue there will be an abstract section in which major research projects, at home or abroad, will be summarized. In the first issue the abstract reports on "Factors Relating to Reading Disabilities in the First Grade of the Elementary School" by Eve Malmquist. The second feature will be a bibliographic section devoted to precise research areas. The first issue contains a selected and annotated bibliography on "The Teaching of Arithmetic" by J. B. Biggs.

The first issue is an attractive publication of eighty pages. It has limp red covers, with the table of contents on the exposed face of the back cover.

This publication should be a welcome accession for the busy educationalist since it will provide extensive and authoritative statements of the research in specific fields, summarized and interpreted by distinguished British scholars. It is edited by W. D. Wall, and published by the Newnes Educational Publishing Company, Limited, once a term, in November, February and June. The price, post free throughout the world, is 5/6 per issue. The annual subscription is 16/6.

G. M. DUNLOP

BOOK REVIEW

WHO ARE THE GUILTY? A study of Education and Crime. David Abrahamsen, M.D., Evergreen E-106. New York, The Grove Press.

The reader's enthusiasm for *Who Are the Guilty?* will depend largely on his psychological orientation, his background in the fields of delinquency and psychology, and his interpretation of the subtitle of the book. Dr. Abrahamsen has written interestingly as a classical Freudian about his work with delinquents who are almost all over eighteen; he draws extensively on anecdotal material from his patients' case histories. He devotes his last two chapters to the role of education; the first is a straight forward intelligent brief for mental hygiene in the schools; the second, a brief for the same approach among parents, children, and teachers. There are several excellent suggestions and a few rather disturbing implications. Dr. Abrahamsen has included an excellent bibliography and an interesting glossary of prison jargon.

If the reader approaches the book in the hope of finding "A Study of Education and Crime", that is, their relationship one to the other, he will be disappointed as was the reviewer. It is almost a handbook of Freudian interpretation of personality and an application to education seems almost an afterthought. Although much of the anecdotal material is interesting, it is often difficult to decide for whom Dr. Abrahamsen is writing. If he is writing for the layman then such references as those to criminals fixated at the oral aggressive or anal sadistic levels for people without sufficient background are inappropriate, as on p. 41.

> In Alex was already laid the foundation for becoming a compulsory individual. He was toilet trained very early and as we saw was also cruel and sadistic, all of which could be led back to a fixation on the anal level . . . When he got himself dirty when he was supposed to be clean, it was a throwback, a regression to his anal stage, his unconscious wish to play with his feces. When he started bedwetting it was unconsciously a substitution for masturbation. He started bedwetting because his mother did not permit him much pleasure while he defecated.

If he is writing for the professional reader, then his detailed explanations of terms like psychosomatic are unnecessary. There are many examples of this sort of ambiguity. For this reader at least the classical Freudian approach with its often too glibly stated professional or lay acceptance, was extremely disturbing, as was the smoothness with which symptoms were slipped into Freudian pigeon holes and then lost. Although Dr. Abrahamsen refers to the ego in his analysis of criminal behavior, his stress, as one would expect from this orientation, is on the inefficacy of the super-ego rather than on the ego controls.

The book wavers between clinical precision, appreciated most by other classical Freudians and those lay people who enjoy snickering at Freudian terminology they do not understand, and trite generalizations such as appears on p. 45.

> We have to emphasize that every factor which prevents the child from a healthy emotional growth will cause either juvenile delinquency and/or mental disturbances, or psychosomatic manifestations.

When a definite stand is taken, the reader is bombarded with a series of qualifications which leave him questioning just what the original statement finally contributed. Anyone involved with work in the field is aware of the many and diverse ramifications involved in personality analysis and any serious attempt to feret out the etiology of criminal behavior. The professional reader will be disturbed that Dr. Abrahamsen's caution does not make for greater clarity and precision.

In the chapters on education Dr. Abrahamsen makes a fervent plea for a more extensive program of mental hygiene in the schools. He does not refer in any detail to projects already in existence. Although he does state the great need for qualified people to work with emotionally disturbed children within the school structure, he does not outline any approach to be used by these people. He discusses the effect of the teacher's personality on the children and seems aware of the strengths and weaknesses to be found in teachers as in any adults. He suggests that if teachers are properly trained (not just another psychology course) and appropriately paid there is an extended role for them in the mental health program. On p. 266 he says:

> Just as the parents' job overlaps the field of learning and culture for their children, so must the teacher be able to recognize emotional difficulties and be able to straighten them out if the development (of the child) is to continue.

To hope that the teacher with some training will be alerted to the difficulties and be able to call proper attention to them is justifiable. But, as is true elsewhere in this book, there is too much ambiguity here. Is he assuming that all parents take on the responsibilities of the job as he has assigned it to them, delinquents' parents included? Does he mean that the teachers' role is the same regardless of the parents' fulfillment? One must not lose sight of the fact that the teacher is primarily a teacher with all the responsibilities and limitations that associate themselves with that role. The teacher is not a therapist, and so long as he or she is a teacher he cannot function in the role of a therapist. How severe are the emotional problems the teacher is expected to "help straighten out"? Abrahamsen has stated several times that the most important periods

of personality development according to Rene Spitz occur in the first year, according to Erickson before the second, and according to Abrahamsen himself, at least before the third. By school age the pattern is deeply set. If the children in question are emotionally disturbed, delinquent, or incipiently delinquent, what is the role of the teacher? He does not speak to these questions. Nor does he refer to the danger of untrained but well meaning people in these roles. The difficulty that arises here stems from the fact that he does not aim this chapter at mental hygiene and delinquency, or the school and delinquency; it is exactly what he has called it "Mental Health and Schools".

In his last chapter he offers suggestions to parents for handling problems. Perhaps these suggestions are all to the good when offered to individuals aware of their inner dynamics and capable of tempering these principles in the light of their own emotional reactions to the situations and the people involved. But this is an extraordinary assumption in view of the limitations he has found in teachers, and in view of the fact that if he does have the parents of deliquents in mind, they themselves are badly confused and in in need of help.

There is too much of the implication that just telling parents and teachers how to handle these problems is sufficient. This is naïve even when dealing with relatively intelligent people who are not emotionally disturbed. These principles must be reached by the individuals themselves and cannot be pronounced as dicta. Nevertheless, these precepts might be tenable in the average (if there is such) classroom and in the family of relatively intelligent parents concerned with the mental health of their children, where emotional distburbance is not overwhelming. He has not told us what we as educators can do for the delinquent.

ANN DOCKRELL

BOOK REVIEW

A STUDY OF INTERPERSONAL RELATIONS—NEW CONTRIBUTIONS TO PSYCHIATRY. Edited with an Introduction by Patrick Mullahy, New York. Grove Press Inc. 1959.

It is a convention among psycho-analysts that teachers' beliefs about psychoanalysis are fixated at an infantile level of direct gratification. Only the earlier and more lurid of Freud's concepts have percolated through to our profession. We have all heard of repressions but very little else.

Of the more recent developments of psycho-analytic thought the ideas of the Ego-psychologists, the group of analysts developing Anna Freud's work and writing in "The Psycho-analytic Study of the Child" have had most impact in Britain and America. They have revolutionized the treatment of disturbed children both in clinics and residential institutions but little of their understanding of the personality and needs of the child has been incorporated in our educational thinking. We read Bettelheim and Redl but have difficulty in developing the controls from within that are the aim of character education.

The other most important development, the psychiatric study of interpersonal relations as practised by the followers of Harry Stack Sullivan, has had no noticeable impact on education at all. Since the Ego-psychologists work largely with children and write books and articles about them we have had to notice their ideas. It is surprising that Educational Psychologists could remain indifferent to a way of thinking that has been so important to Sociologists and Anthropologists as Sullivan's. Educators too are concerned largely with interpersonal-relations, our relations with our pupils are our primary educational medium.

This book is not a systematic exposition of Sullivan's views. It contains a few fundamental articles and also a number of applications of these ideas to specific situations. Some of the articles are of interest primarily to practising therapists, for example Bruch's article on obesity and Thomson's on homosexuality. Readers familiar with conventional psycho-analytic writing on these topics cannot but be impressed by this approach. The rigid categorizing of some primitive Freudians, the insistence that each behaviour syndrome can be linked to a failure in a specific point of infantile development are rejected. Instead individual variations in environmental experiences are given full weight. Many of the insights of the conventional psycho-analysts are used but there is concern about the functioning role of a neurosis at this point in time. A neurotic behaviour pattern, or as Sullivan prefers to say a parataxic dynamism, is not merely a relic of some earlier experience but has a role in the adjustment of the personality to current situations. Sullivan and his followers are interested in the history of a particular syndrome but prefer to treat the existing personality rather than to reconstruct past experiences.

The articles in which the basic tenets of the system are expounded are of interest to all working in the field of personality. The rapid transition in Sullivan from the mystical, "empathy" between mother and child, to the physiological, the role of the striped skeletal muscles in relieving anxiety, together with much unsub-

stantiated speculation alienates the psychologist. When Mullahy says that "noxious emotional states" experienced by a baby when a mother manifests hostility to him are "one of the most terrible experiences man undergoes . . . somewhat like a blow on the head" we want to ask how he knows and in what ways they are "like a blow on the head". Nevertheless Sullivan and Mullahy can contribute a great deal to our understanding of personality maladjustments and their disturbance of the perceptual field. Psychologists and educators interested in the work of Dewey and in phenomenological psychology will be struck by the similarity of basic concepts. There is a possibility here of a more fruitful reconciliation between academic psychology and psycho-analysis than we have seen so far.

Sullivan's language is sometimes obscure and his subjects of study are far removed from the classroom but the articles by Talcott Parsons and Ruth Benedict are fundamental to our thinking about Mental Hygiene in school. Though they present findings of earlier researches they are pertinent to our present educational practices. Through the writings of Allison Davis and others we are familiar with the concept of sociolized anxiety. Parsons' lucid and pregnant analysis of the sources of aggression in our social structure can illuminate the disastrous consequences of much of our pedagogical practice without the comforting illusion that we are really storing energy for future use. The anger and aggression in our children that we adults find so threatening may be a manifestation of an instinctual death urge but it is much more fruitful to look to our child-rearing and classroom practices that generate this anger and prohibit its expression so that it is turned inward to become ulcers and neurasthenia or outward in some socially acceptable form of nigger-baiting and McCarthysim.

Benedict makes a disturbing analysis of our practices in three areas, responsible-non-responsible status roles, dominance-submission, and sex roles. It isn't merely that in our society we expect submission in our boys and dominance in our men but our "cultural institutions furnish no adequate support to the individuals as he progresses from role to role." In the army a private soldier's duty is to obey, a commanding officer's is to command. In between are officers and N.C.O.'s whose role combines dominance and submission in varied degrees. Most soldiers move from complete submission to some degree of dominance, but each step, each increment in dominance and corresponding decline in submission, is marked by a change in rank and the attitude of equals, inferiors, and superiors is dictated largely by status symbols, chevrons or pips. There is a carefully graded hierarchy permitting gradual precisely

defined changes in status role. In our more open society there is no system of progression to support the transition of the adolescent from one role to another. Frosh weeks are a poor substitute for the elaborate hierarchies of fag, middle school and prefect with which readers of "Tom Brown's Schooldays" are familiar. Benedict goes on to say that "The adolescent period of Sturm und Drang with which we are so familiar becomes intelligible in terms of our discontinous social institutions." Adolescence is so protracted in our society that there is no accepted transition. "Continuity in sex expression means that the child is taught nothing it must unlearn later . . . The adult in our culture has often failed to unlearn the wickedness or dangerousness of sex, a lesson which was impressed upon him strongly in his formative years." How many of these lessons that have to be unlearned later are learned in school?

W. B. Dockrell

Composite High Schools in Canada

University of Alberta Monographs in Education No. 1

- Reports a study conducted by the Division of Educational Administration on the status of the Composite High Schools in the various provinces of Canada.
- Contains chapters contributed by authorities on secondary education from the different provinces.
- Draws conclusions about present status and makes recommendations for the future development of the Composite High Schools.
- Edited by John H. M. Andrews and Alan F. Brown.
- Cost—one dollar.

Copies may be obtained from:

The Committee on Educational Research
Faculty of Education
University of Alberta
Edmonton, Alberta